Enzymatic Ring Formation in Polyether Tetronate Antibiotic Biosynthesis

Rory Fox Little

Trinity College

Department of Biochemistry

University of Cambridge

This dissertation is submitted for the degree of Doctor of Philosophy

May 2019

Preface

This dissertation is the result of my own work and includes nothing that is the outcome of experiments done in collaboration except where specifically indicated in the text. The work described in this dissertation was carried out under the supervision of Professor Peter F. Leadlay in the Department of Biochemistry of the University of Cambridge, United Kingdom, between October 2015 and May 2019. It is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution. I further state that no substantial part of my thesis has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University or similar institution. It does not exceed the limit of 60,000 words set by the Biology Degree Committee.

Rory Fox Little

Trinity College
University of Cambridge
May 2019

Journal publication arising from this thesis:

Little, R,. Paiva, F., Jenkins, R., Hong, H., Yuhui, S,. Demydchuk, Y,. Samborskyy, M,. Tosin, M,. Leeper, F,. Dias,. M,. Leadlay, P,. Unexpected Enzyme-Catalysed [4+2] Cycloaddition and Rearrangement in Polyether Antibiotic Biosynthesis. *Nature Catalysis* (2019). DOI: 10.1038/s41929-019-0351-2.

Abstract

Enzymatic Ring Formation in Polyether Tetronate Antibiotic Biosynthesis

Rory Fox Little

The formation of rings in carbon backbones is essential for the biological activity of many natural products. The polyether tetronate antibiotics tetronasin, tetronomycin, and tetromadurin (SF2487/A80577) are notable for their ring diversity, each possessing a tetronate, cyclohexane, tetrahydropyran, and at least one tetrahydrofuran ring. The antibiotic activity and complexity of these polyether tetronates has led to research on their biosynthesis from actinomycete bacteria. Despite this, the mechanism of stereospecific cyclohexane and tetrahydropyran formation has remained mysterious. Although no formal [4+2] cycloaddition (Diels-Alder reaction) is predicted in the biosynthesis of these compounds, the biosynthetic gene clusters of all three were found to contain a pair of genes that encode homologues of two different [4+2] cyclases previously identified in complex spirotetronate pathways. Specific gene deletions demonstrated that both classes of [4+2] cyclase homologue are essential for polyether tetronate biosynthesis. In the tetronasin producer, *Streptomyces longisporoflavus,* deletion of the [4+2] cyclase enzyme homologue gene *tsn11* resulted in production of a new metabolite that was characterised as an open form of tetronasin lacking both cyclohexane and tetrahydropyran rings. Incubating this metabolite with purified Tsn11 resulted in the production of an unknown intermediate labelled T-22. The structure of T-22 was determined using NMR to contain an unexpected oxadecalin moiety but still lack the tetrahydropyran ring, implicating Tsn11 as catalysing an apparent inverse-electron-demand hetero-Diels-Alder reaction. Remarkably, incubating T-22 with purified Tsn15, the other [4+2] cyclase homologue, formed the tetrahydropyran ring and fragmented the oxadecalin moiety to a cyclohexane ring, producing tetronasin. To gain structural insight into the novel activity of Tsn15 it was successfully crystallised. The structure of Tsn15 was then solved at 1.8 Å using SAD phasing; and Brazilian collaborators solved a Tsn15-ligand structure at 1.7 Å. Tsn15 shares the same eight-stranded β-barrel fold as its [4+2] cyclase homologues. The two main mechanisms considered here for the Tsn15-catalysed conversion of T-22 into tetronasin were a general acid/ base or a pericyclic mechanism. Site-directed mutagenesis of the Tsn15 active site indicated that none of the acid/base amino acid side chains were essential for activity, favouring instead the pericyclic mechanism.

Acknowledgements

First and foremost, I am extremely grateful to my supervisor Professor Peter Leadlay. I am honored to have had the opportunity to develop as a scientist in your laboratory. Thank you for taking a chance on a student from the other side of world. I have learnt so much from you about polyketides, biosynthesis, and science in general, that I will carry into the future.

I am grateful to and will miss all of the wonderful members of the Leadlay lab I have worked with. In no particular order, thank you to Dr. Hui Hong, Dr. Katsiaryna Usachova, Dr. Oksana Bilyk, Dr. Karen Chan, Dr. Fanglu Huang, Dr. Freddie Dudbridge, Dr. Katharina Dornblut, Dr. Jacob Pollock, Dr. Kwaku Kyeremeh, Dr. Annabel Murphy, Dr. Marie Yurkovich, Dr. Anna Reva, Oana Sadiq, Felix Trottmann, Roel van Harten, Carten Schotte, Johanna Ritzer, Sebastian Walesch, Shilo Dickens, Anna Efimova, Natalia Scott, Reda Deglau, Hannah Büttner, and Vinzent Schultz. Special thanks to everyone from the third floor ("upstairs") lab for creating a fantastic working environment and putting up with my hundreds of agar plates. Thanks to the whole DNA sequencing team for your excellent work. A special mention to Dr. Hui Hong and Dr. Fanglu Huang, two of the most inspiring people I have ever met. Thank you for everything you taught me.

I would like to thank all of the collaborators who contributed to this project. Thank you to Fernanda Paiva and Dr. Marcio Dias for your protein crystallisation magic. A huge thanks to Dr. Finian Leeper for all of your help of the NMR assignments and for patiently answering my (many) questions on reactions mechanisms. Finally, thank you to the Dr. Manuela Tosin and Robert Jenkins for performing the polyketide offloading experiments in Chapter 5. Thanks to Dr. Markiyan Samborskyy for assembling the *A. verrucosispora* genome. Thanks to Dr. Bill Broadhurst, Professor Ernest Laue, and Professor Gregory Challis for their discussions on the project. I am also grateful to Dr. Dima Chirgadze for crystallography assistance and Dr. Katherine Stott of the Biophysics facility.

Thank you to my parents and brother for their constant support and encouragement. A big thank you to all of my friends for getting me out of the lab to do fun things. I am grateful for the support and opportunities provided by Trinity College. I also cannot express enough gratitude to the Woolf Fisher Trust and the Cambridge Commonwealth Trust for funding me; it was great to have Nigel and Jan visit each year. Finally, to show last is certainly not least, thank you to Olivia Tidswell for your endless love, support, and sweetness.

Table of Contents

Abbreviations

2D	two-dimensional
3D	three-dimensional
A_{600}	absorbance at 600 nm
ACP	acyl carrier protein
AT	acyltransferase
Apr	apramycin
BGC	biosynthetic gene cluster
BLAST	Basic Local Alignment Search Tool
BSA	bovine serum albumin
bp	base pair
COSY	correlation spectroscopy
Da	Dalton
DEBS	6-deoxyerythronolide B synthase
DH	dehydratase
DNA	deoxyribonucleic acid
EDG	electron withdrawing group
EDTA	ethylenediaminetetraacetic acid
ER	enoylreductase
ESI	electrospray ionisation
EWG	electron donating group
gDNA	genomic DNA
HEPES	N-(2-hydroxyethyl)piperazine-N'-(2-ethanesulfonic acid)
HMBC	heteronuclear multiple bond correlation
HOMO	highest occupied molecular orbital
HPLC	high performance liquid chromatography
HRMS	high resolution mass spectrometry
HSQC	heteronuclear single quantum correlation
$_{INV}$HDA	inverse-electron-demand hetero-Diels-Alder
IPTG	isopropyl-β-D-thiogalactopyranoside
kbp	kilo base pairs
KR	ketoreductase
KS	ketosynthase
LUMO	lowest unoccupied molecular orbital
mad	tetromadurin
Mbp	Mega base pairs
MES	2-(N-morpholino)ethanesulfonic acid
MS	mass spectrometry
m/z	mass to charge ratio
NADH	nicotinamide adenine dinucleotide, reduced form
NADPH	nicotinamide adenine dinucleotide phosphate, reduced form
NOESY	nuclear Overhauser effect spectroscopy
orf	open reading frame
PCR	polymerase chain reaction

PDA	photodiode array
PDB	Protein Data Bank
PEG	polyethylene glycol
Phyre2	protein homology/analogy recognition engine 2
PKS	polyketide synthase
PPTase	4'-phosphopantetheinyl transferase
rpm	revolutions per minute
SDS	sodium dodecyl sulphate
sp., spp.	species
SPE	solid phase extraction
$S_{20,w}$	sedimentation coefficient corrected for the density and viscosity of water at 20 °C
T_m	annealing temperature
TE	thioesterase
TFA	trifluoroacetic acid
TIC	total ion current
tmn	tetronomycin
TOCSY	total correlated spectroscopy
tsn	tetronasin
TSB	tryptic soy broth
UV	ultra violet
v/v	volume to volume ratio
WT	wild type

List of Figures and Tables

Chapter 1: Introduction

1.1 Bacterial natural products

1.1.1 Nature as a source of medicines

Nature is an essential source for human medicines. Written records dating back to ancient Mesopotamia and Egypt both detail the use of plant oils to treat aliments such as coughs, inflammation, and infections[1]. Traditional Chinese medicine is also largely based on deriving medicines from natural sources such as plants and animals. A notable example is the herb *Artemisia annua*, which was used for centuries in China to treat malaria. In 1971, the compound artemisinin was isolated from *A. annua* and shown to be the active anti-malarial component[2]. Since then, artemisinin and derivatives have been essential drugs for controlling malaria, recognised by awarding the key scientist involved in its discovery, Dr. Tu Youyou, the 2015 Nobel Prize in Medicine[3].

Small molecule compounds such as artemisinin that are isolated from natural sources are collectively called natural products (**Figure 1.1**). Natural products are also typically classified as secondary metabolites: metabolites that, in contrast to primary metabolites, are not essential for the growth of the organism but may provide a selective advantage under certain circumstances[4]. Aside from plants, medically important natural products are also derived from microorganisms, such as fungi and (particularly) bacteria. Natural products and their derivatives are significantly represented in the list of drugs approved by US Food and Drug Administration (FDA)[5]. As of 2014, 33% of all FDA approved drugs were natural products or their derivatives[5]. The percentage increases to 54% if synthetic compounds that mimic natural product structures are also included[5]. Natural products make good drug candidates because, in addition to possessing desirable characteristics such as high levels of structural and stereochemical diversity, they are by definition biologically relevant molecules that have evolved to interact with living systems[6]. Despite this, an issue with developing any natural product drug is the problem of supply. From a given biological source (such as a plant, bacteria, fungus etc…) a natural product of interest may represent only 1% of the total crude extract by mass[7]. In some cases it is possible to increase yields by scaling up cultivation, but many producers (such as a

bacterium symbiotically associated with a deep sea sponge) are not amenable to being farmed/cultivated on a large scale.

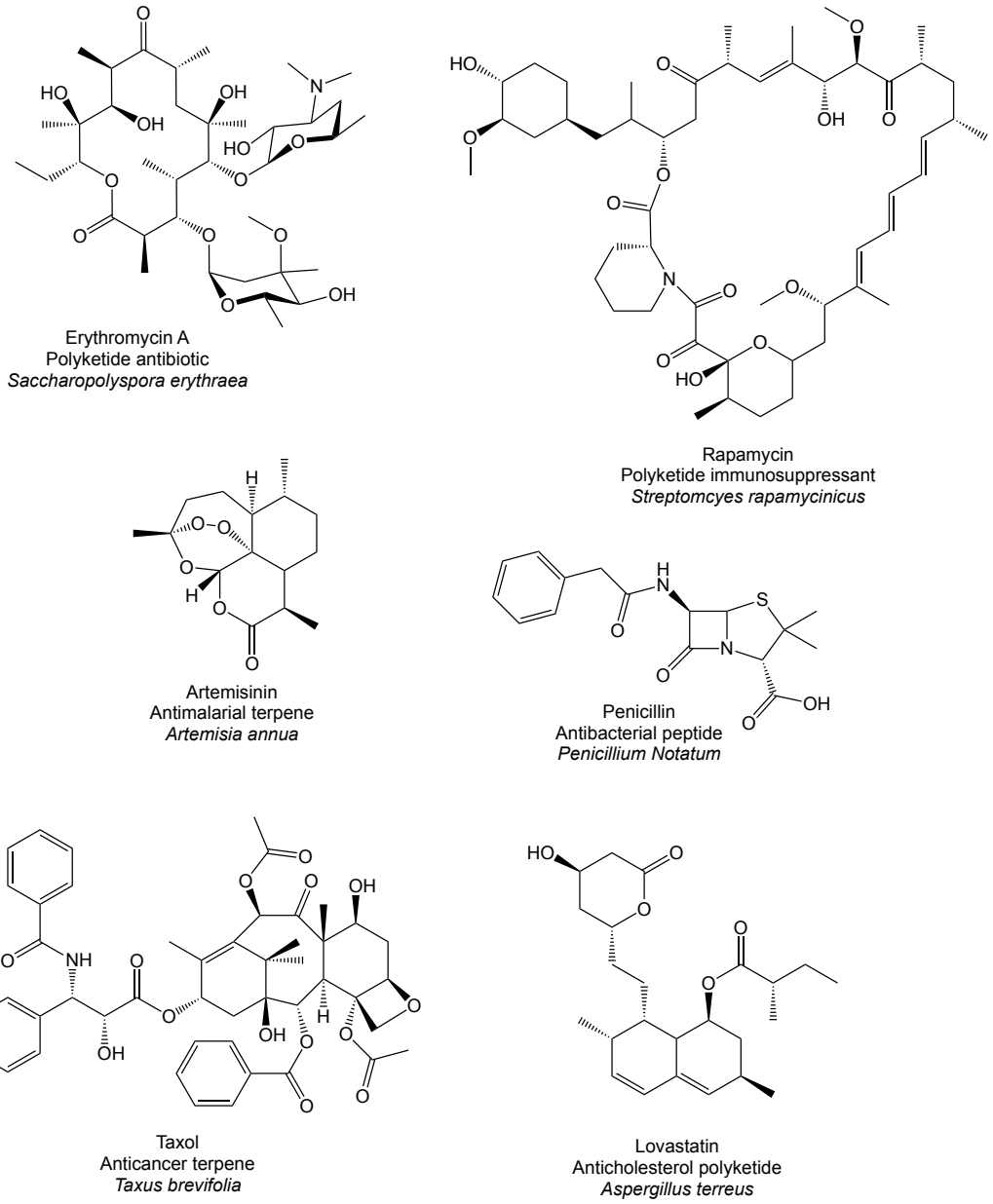


Figure 1.1 Structures of prominent. medically important natural products.

An alternative strategy to produce large amounts of a natural product is *via* total chemical synthesis. However, the complex structures of natural products, often containing multiple chiral centres, may render a wholly synthetic approach too difficult or costly for the required scale[7,8]. Chemical semisynthesis has proved a versatile option, where the parent natural

product is readily available. Such work has had most impact on the development of anti-infective molecules, and natural-product based entities represent 73% of all FDA-approved antibiotics[5]. A notable example of an antibiotic natural product is penicillin, isolated in 1928 from the *Penicillium notatum* fungus[9]. Penicillin came to widespread public attention for its role in treating infections during World War II[10]. Despite its initial success, patients containing penicillin-resistant strains of *Staphylococcus aureus* bacteria were detected just over a decade later in 1942[11]. Since then, the development of antibiotic resistance in bacterial pathogens has only become more prevalent and is a major public health problem today.

1.1.2 Antibiotic resistance

Shortly after the discovery of penicillin resistance, what is described as the "golden" era of antibiotic discovery began. Between 1940 and 1970 almost all of the major classes of antibiotic, including the tetracyclines, macrolides, and aminoglycosides, were discovered[12,13]. During the same period, antibiotic resistance to these new drugs also developed, owing to the selection pressure the antibiotics placed on the bacterial pathogens. A bacterial pathogen that, by chance, acquires a mutation that confers resistance to a given antibiotic is selected for and can generate a resistant population. Genes that confer antibiotic resistance (resistance genes) can even be transferred from a resistant strain to a non-resistant strain by a process called horizontal gene transfer[14]. Numerous mechanisms of antibiotic resistance in bacteria have been characterised[15]. These include:

1) Reduced membrane permeability. Hydrophilic antibiotics can only cross a bacterial plasma membrane though transmembrane protein channels. Down-regulation of these protein channels can therefore decrease entry of the antibiotic into the cell.

2) Upregulation of antibiotic efflux. Bacteria contain transmembrane efflux proteins that often have broad substrate selectivity, enabling the cell to export antibiotics from the cytoplasm. The overexpression of these efflux pumps can therefore enhance antibiotic resistance.

3) Mutation/alteration of the antibiotic target. The change in the antibiotic target need not be through a mutation, as chemical modification of the target (such as methylation or acetylation) can also interfere with antibiotic binding.

4) Antibiotic inactivation. Enzymes within the bacterial pathogen may hydrolyse the antibiotic or catalyse a chemical modification of its structure that renders it inactive.

Taken together, these resistance mechanisms have led to the development of multidrug-resistant and even totally drug-resistant bacterial pathogens[15,16]. Combating antibiotic resistance is a multi-faceted problem requiring political will, such as implementing policies that will prevent antibiotic over/misuse, as much as developing replacement antibiotics. From 1990 onwards, advances in DNA sequencing and genetic engineering technology have provided an unprecedented opportunity to study how natural products are produced by nature i.e., natural product biosynthesis. Understanding the mechanisms of natural product biosynthesis has the potential for antibiotic-producing organisms to be genetically manipulated to increase product yields or even produce novel analogues. As bacteria are some of the most prolific producers of antibiotics and other natural products, most of the research on natural product biosynthesis has been conducted on them.

1.1.3 *Streptomyces* bacteria

Bacteria of the Gram-positive actinobacterial phylum (actinomycetes) produce two-thirds of all known antibiotic compounds, with the majority contributed by the *Streptomyces* genus (streptomycetes)[17]. The streptomycetes likely evolved around 450 million years ago in response to the colonisation of land by green plants providing a new nutrient source[18]. They typically live in soil and are saprophytic (consume decaying organic material). The breakdown of plant material by streptomycetes is achieved though the extracellular secretion of catabolic enzymes, such as cellulases[18]. In the case of *Streptomyces coelicolor*, upwards of 800 different proteins are predicted to be exported to utilise decaying plant and soil nutrients[19]. Streptomycetes clearly compete with fungi for resources, as many secrete chitinases to digest fungal cell walls[20] or produce antifungal natural products[21–23]. In addition to a soil environment, streptomycetes are also found in the ocean[24,25], including the Arctic[26], and in symbiotic relationships with eukaryotic organisms. Streptomycetes have been found living on wasps[27,28], ants[21,29,30], beetles[31], fungi[32], and plants[33,34], often producing antibiotic or antifungal compounds that protect their host. Other interactions with plants are pathogenic, notably potato scab caused by *Streptomyces scabies*[35]. In rare cases streptomycetes also infect humans to cause actinomycetoma[36,37].

The phylum name actinomycete is derived from the Greek words “actis”, meaning ray or beam, and “mykes” meaning fungus. The actinomycetes were given a name meaning “ray fungus” because they were originally misidentified as such[18]. Such a mistake was understandable, as under a microscope actinomycetes often resemble a branching fungal mycelium. The fungal comparison is even stronger in the streptomycete genus, as many of its species reproduce by sporulation, as fungi do. A typical streptomycete life cycle starts with a mononucleate spore germinating and forming a polynucleate, often pigmented, substrate mycelium beneath it, releasing extracellular enzymes to solubilise essential nutrients for itself from its surroundings[17]. Antibiotics are often secreted at the same time to protect the solubilised nutrients from competitor microorganisms[18]. In response to depletion of nutrients, the substrate mycelium is programmed to lyse, supplying a burst of nutrients for aerial hyphae to form and differentiate into white, mononucleate, hydrophobic spores for dispersal[17,18] (**Figure 1.2**). This transition typically is accompanied by an increase in the number and level of antibiotics released. Most *Streptomyces* species are present in soil in the spore form[38], only germinating when conditions are favourable[39]. In contrast, other streptomycetes contain mutations in the *bldA* “bald” locus and do not form spores[39].

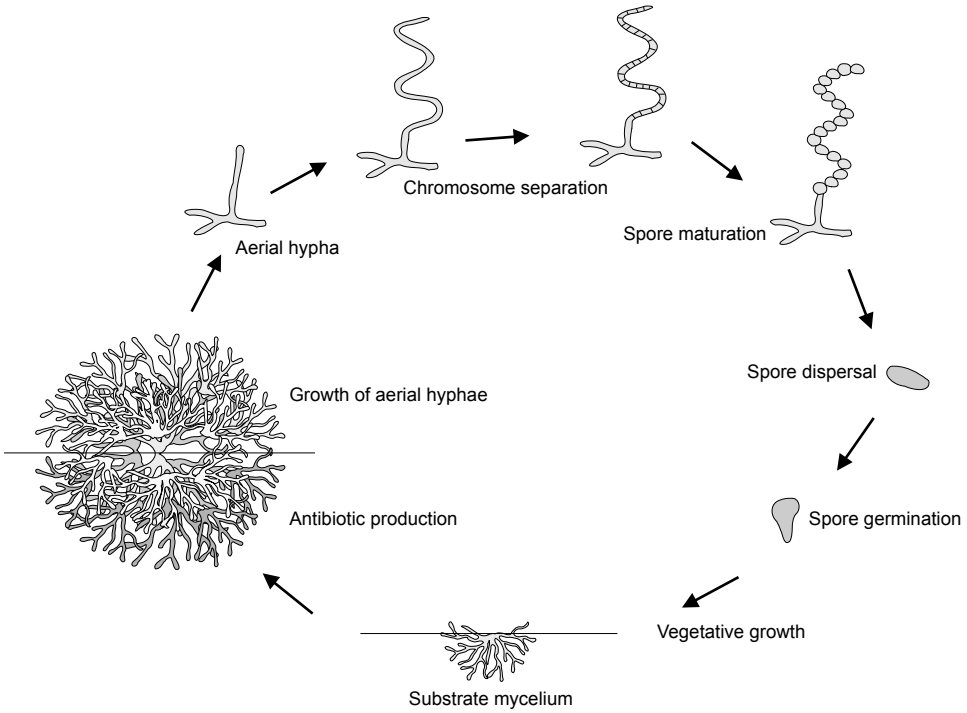


Figure 1.2 Life Cycle of a Sporulating Streptomycete. Following germination from a single spore, the streptomycete life cycle begins with vegetative growth to form a substrate mycelium. Aerial hyphae then begin to grow and differentiate into mononucleate spores for reproductive dispersal. Antibiotics are often produced during growth of the aerial hyphae to protect the soluble nutrients produced by the catabolic enzymes secreted by the streptomycete. Figure adapted from[17].

1.2 Natural product biosynthesis

1.2.1 Biosynthetic gene clusters

Streptomycetes produce a range of natural product classes including polyketides (which will be discussed in greater detail later), non-ribosomal peptides, terpenes, and ribosomally synthesised and post-translationally modified peptides (RiPPs). The genes responsible for the biosynthesis of these natural products are encoded in the bacterial genome or on accessory plasmids. Due to the complexity of natural products, multiple genes are often required for their biosynthesis. The genes responsible for the biosynthesis of a given natural product are also typically grouped together in a biosynthetic gene cluster (BGC). BGCs can reach over 200 kbp in size[40] and can contain more than 40 genes, all of which contribute to the biosynthesis of a single natural product. Genes may directly contribute to the construction of the natural product or have a role in resistance or transcriptional regulation. A streptomycete genome is comprised of a single linear chromosome[41] with BGCs mainly concentrated near the chromosome ends[19,42,43]. The number of BGCs within a bacterial species increases with its genome size, with roughly 2.4 BGC per Mbp, resulting in an average of 3.7% of the genome being BGCs[44]. Exceptions exist, notably *Streptomyces bingchenggensis* in which 22%[44] of its 12 Mbp[43] genome encodes BGCs for natural product biosynthesis (2.6 Mbp in total). The genes in a BGC are often organised into operons, enabling coregulation and cotranscription of multiple genes. BGCs also typically contain genes encoding regulatory elements that control transcription of the entire cluster.

The reason why genes involved in natural product biosynthesis evolved to cluster with one another has been the subject of much debate. The prevailing theory is called the “selfish operon” model, and proposes that biosynthetic genes cluster to increase their chance of it being passed on via horizontal gene transfer[45–47]. As maintaining a large BGC is metabolically expensive, unless the natural product they produce provides a survival advantage BGCs are typically eliminated from the genome. Dispersal via horizontal gene transfer can therefore save a BGC from extinction if does not provide a reproductive advantage in its current bacterial host. By having all of the biosynthetic genes clustered together, the complete cluster has a better chance of being passed on intact to a different host, where the encoded natural product does provide a reproductive advantage[45].

Phylogenetic analysis of bacterial operons supports this model, indicating that horizontal transfer of entire operons is the main process creating co-localised foreign genes in bacterial genomes[47]. A piece of evidence against the selfish operon model is that the essential bacterial genes (such as those in primary metabolism) with related functions also cluster together and form operons[48]. However, it is likely that essential genes and secondary metabolite genes cluster for different reasons[49]. Essential genes may cluster together in coregulated operons to decrease the chances of a gene deletion disrupting one of them[49]. In contrast, if essential genes are uniformly distributed in the genome and independently regulated, the chances that a gene deletion will affect one of them, and thereby dismantle the entire metabolic process, is higher[49]. Regardless of why it occurs, the clustering of biosynthetic genes has simplified the identification and study of the BGCs responsible for natural product biosynthesis. The polyketide natural products are some of most valuable and their biosynthesis has been the most extensively studied.

1.2.2 Polyketide natural products

Polyketides are a structurally diverse and medically important family of natural product. Polyketide carbon skeletons are mainly constructed from the successive condensation of acetate and propionate units, each derived from an activated form of malonyl-CoA (co-enzyme A) or (*2S*)-methylmalonyl-CoA, respectively (**Figure 1.3**).

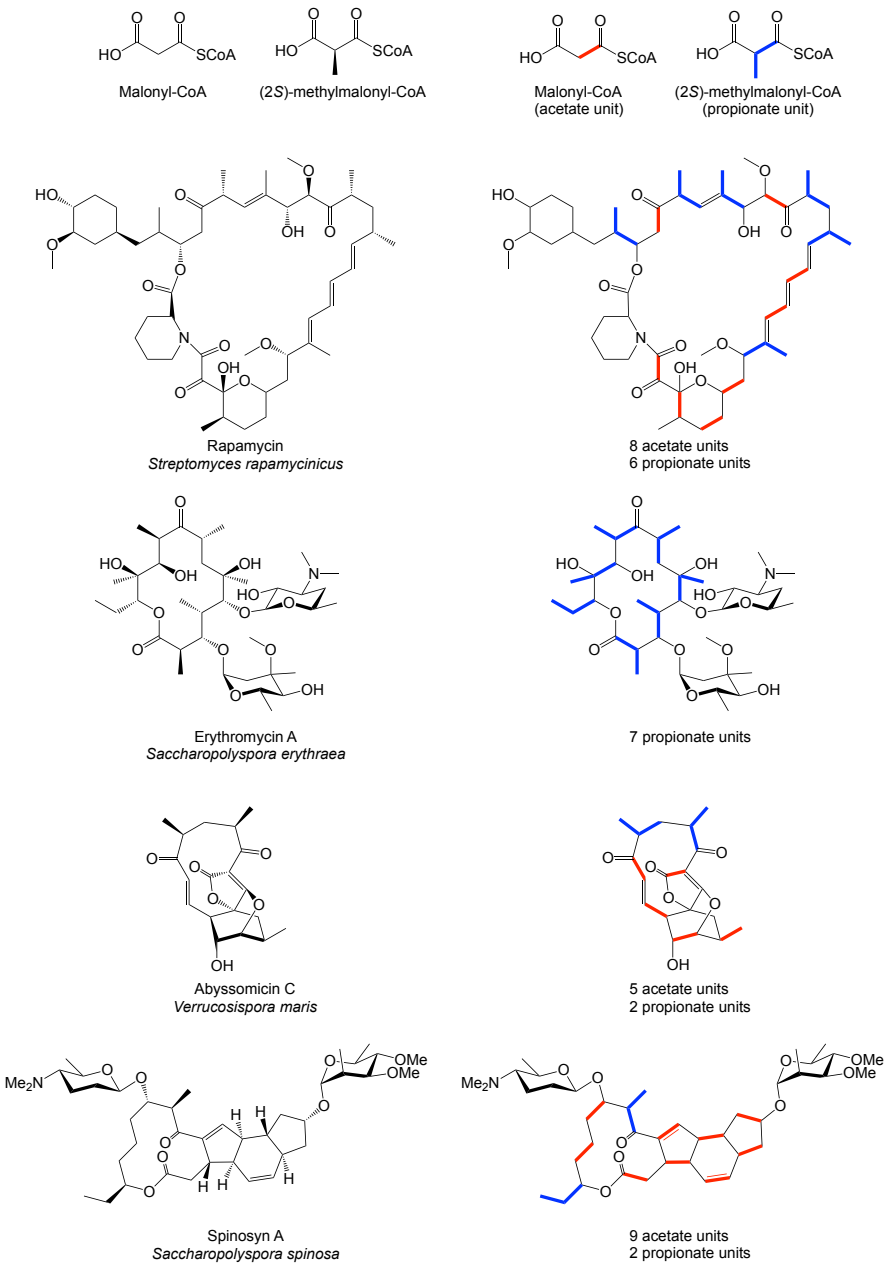


Figure 1.3 Acetate and propionate units in polyketide carbon skeletons. The carbon skeletons of polyketides typically consist of acetate and propionate units derived from malonyl-CoA and (*2S*)-methylmalonyl-CoA, respectively. The structures of four polyketide natural products are shown with the two-carbon acetate units shown in red, and the three-carbon propionate units shown in blue.

In addition to actinobacteria, polyketides are produced by other bacteria, and also by plants[50] and fungi[51]. A recent study even unexpectedly revealed a polyketide produced by a nematode[52]. Polyketides broadly comprise two different classes: the aromatic polyketides and the reduced polyketides. Aromatic polyketides contain fused polyphenol rings derived from unreduced polyketide chains. In contrast, reduced polyketides contain multiple reduced carbonyl functional groups and rarely contain aromatic moieties[53]. Polyketides are

one of the most clinically successful natural products classes, with roughly 1% of all known polyketides possessing a drug-like activity, five times higher than the natural product average[54]. Sales of polyketide drugs annually exceed 20 billion USD[55]. Clinically used polyketide drugs display a wide range of activities, ranging from antibiotic (erythromycin A, rifamycin S), immunosuppressant (rapamycin), anticholesterol (lovastatin) or antifungal (amphotericin B) to anticancer (epothiolone, doxorubicin, calicheamicin)[53,56,57]. The potent insecticide spinosyn A and its derivatives are also polyketides[58]. In several notable cases, other polyketides are causative agents in disease, such as mycolactone A (produced by *Mycobacterium ulcerans*) in Buruli ulcer formation[59], and brevetoxin B (produced by the dinoflagellate *Karenia brevis*), responsible for neurotoxic shellfish poisoning[59,60] (**Figure 1.4**).

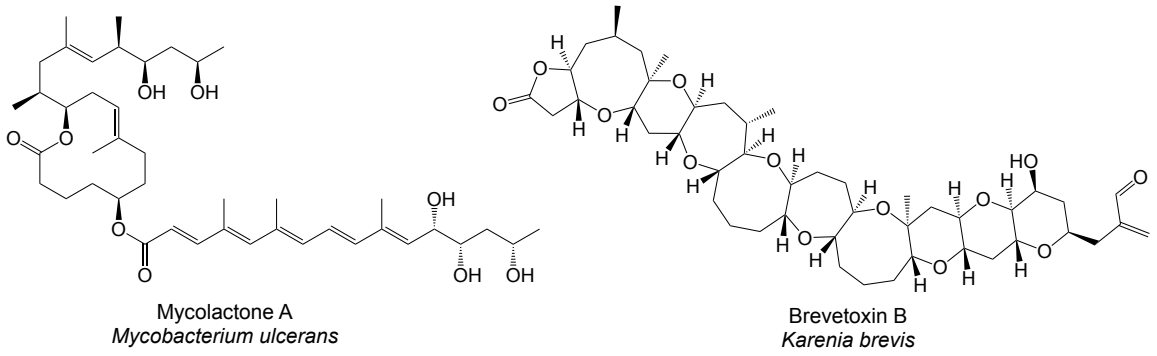


Figure 1.4 Structures of the pathogenic polyketides mycolactone A and brevetoxin B.

1.2.3 Polyketide biosynthesis

Polyketide natural products are produced by polyketide synthase enzymes (PKSs). There are three different classes of PKS: type I, type II, and type III. Polyketides produced by each type of PKS are accordingly called a type I, II, or III polyketide. While malonyl-CoA and (*2S*)-methylmalonyl-CoA are the most common “building blocks” (also called extension units) utilised by PKS enzymes, others, such as (*2S*)-ethylmalonyl-CoA[61], are sometimes used. The chemical properties of thioesters of co-enzyme A facilitate the covalent transfer of the extension unit to the PKS enzyme, where they can be incorporated into the growing polyketide acyl chain. Once attached to the PKS via the acyl carrier protein (ACP) domain, each extender unit undergoes a Claisen-like decarboxylative condensation reaction together with the nascent polyketide chain tethered to the ketosynthase (KS)

enzyme/domain[57]. This contributes two carbons to the polyketide carbon skeleton in the case of malonyl-CoA, and if the extension unit is (2*S*)-methylmalonyl-CoA, also an α-methyl branch[57]. Successive chain elongation condensation reactions produce the polyketide carbon backbone. In this sense, the process of polyketide biosynthesis is highly reminiscent of fatty acid biosynthesis[62]. However, in contrast to fatty acid biosynthesis, where each newly-introduced β-ketothioester carbonyl group is fully reduced to form a saturated acyl chain, in polyketide biosynthesis the degree of carbonyl reduction is variable, introducing chemical variation along the polyketide chain and leading to greater structural diversity. Once the final polyketide chain length is reached it is released from the PKS, often in a cyclised form. Additional enzymes, called post-PKS enzymes, expressed from the BGC may then further tailor the structure of the polyketide, adding chemical groups such as sugars and hydroxyls or catalysing additional cyclisation reactions.

Type II PKSs are responsible for the biosynthesis of aromatic polyketide natural products such as the medically important tetracycline antibiotics[63]. Unlike the modular type I PKSs, where multiple catalytic domains are required for a single round of chain extension (discussed in the next section), type II PKSs consist of iteratively acting standalone proteins[63]. A minimal type II PKS biosynthesis pathway consists of a standalone acyl carrier protein and two proteins that resemble ketosynthase domains: a KS_α and KS_β. These two KS proteins are highly similar to one another, and together catalyse the necessary chain initiation and extension events required for polyketide biosynthesis[63]. Like type II PKSs, type III PKSs also produce aromatic products through the iterative condensation of malonyl-CoA onto (for example) a cinnamoyl-CoA starter molecule. Present in plants, fungi, and bacteria, type III PKSs consist of a single KS-like protein and are typically ACP-independent[64]. The main focus of this thesis, however, is on the function of type I polyketide synthases.

1.2.4 Type I polyketide synthases

Modular type I PKSs are large (typically >100 kDa) homodimeric enzymes responsible for the biosynthesis of many important polyketide natural products, including macrolides and polyethers[55]. The structure of these massive modular complexes is an on going and difficult area of research, covered well in several recent reviews[65,66]. Several levels of organisation

exist within type I PKSs. Each type I PKS enzyme consists of at least one, though usually multiple, modules. Each module is responsible for the selection and incorporation of an extension unit (typically malonyl-CoA or (2*S*)-methylmalonyl-CoA) into the growing polyketide chain. The number of modules present in a type I PKS therefore dictates (and can be used to predict) the length of the polyketide product it produces. There is a separate class of single-module type I PKS common in fungi where the module acts iteratively to build up the chain, with optional input from additional stand-alone enzymes[67].

Type I PKS modules are, in turn, comprised of discrete regions of catalytic activity called domains. There are seven types of catalytic domain typically found in type I PKSs: the essential core domains are the acyltransferase (AT), ketosynthase (KS) and an acyl carrier protein (ACP) domain bearing a 4'-phosphopantetheine prosthetic group. Additional domains that may be present in a module include ketoreductase (KR), dehydratase (DH), enoylreductase (ER), and a chain-releasing domain, often a thioesterase (TE) domain. Other domains have been identified in several polyketide biosynthesis pathways, such as domains that catalyse pyran ring formation (PS domains)[68,69], sulphur incorporation[70], and polyketide chain branching[71], but these are less common. To catalyse a specific round of polyketide chain extension each module requires the core AT, ACP, and KS domains. The role of the AT domain is to select the extension unit and attach it the adjacent ACP domain via the 4'-phosphopantethine linker. In some modular type I polyketide pathways, called *trans*-AT pathways, the AT domains are provided as standalone proteins[72]. An example of a polyketide produced by a *trans*-AT PKS is the polyether antibiotic pederin, produced by an unculturable *Pseudomonas* sp. found on Paederus beetles[73]. The KS domain then catalyses carbon-carbon bond formation to incorporate the extension unit into the growing polyketide acyl chain. Any reduction domains present in the module can then modify the polyketide acyl chain, reducing the β-keto group to form a secondary alcohol (if only a KR domain is present), an *E*-double bond (if both a KR and DH is present), or saturated acyl chain (if a KR, DH, and ER domain is present). The acyl chain is then passed on to the KS domain of the neighbouring module, where the process repeats. Final release of the full-length polyketide from the terminal module is typically accompanied by cyclisation. The biochemistry and mechanisms of different domain types will be discussed in more detail in **Section 1.3**.

The first modular type I PKS genes to be sequenced were those from the BGC of

erythromycin A in *Saccharopolyspora erythraea*[74–77]. The biosynthesis of erythromycin A is a good example to illustrate the key process and mechanisms of type I PKSs. Erythromycin A is a 14-membered macrolide antibiotic that inhibits translation by binding to the 50S subunit of the prokaryote ribosome[78,79]. Three genes, *eryAI*, *eryAII*, and *eryAIII*, respectively encode the type I PKS enzymes DEBS1 (370 kDa), DEBS2 (374 kDa), and DEBS3 (333 kDa). Together the three homodimeric DEBS proteins form a complex close to 2 MDa in size, roughly the same size as the prokaryote ribosome erythromycin A inhibits[65]. DEBS1, DEBS2, and DEBS3 work together to produce 6-deoxyerythronolide-B, the macrocyclic precursor of erythromycin A, from seven propionate units. Collectively the three DEBS proteins contain seven modules that act in a specific order as a molecular "assembly line" to correctly synthesise 6-deoxyerythronolide-B. DEBS1 initiates 6-deoxyerythronolide-B biosynthesis, and consists of three modules: a loading module and two extension modules (**Figure 1.5**). The loading module consists of an AT and ACP domain, while extension modules 1 and 2 additionally both contain a KS and KR domain. In the case of DEBS1, the AT of the loading module actually selects propionyl-CoA rather than (2*S*)-methylmalonyl-CoA, and attaches it to the ACP domain of the loading module. In most type I PKS pathways, the loading module contains a special KS domain called a KS^Q domain that, while unable to catalyse chain extension, can catalyse the decarboxylation of (2*S*)-methylmalonyl-CoA to propionyl-CoA once it is attached to the ACP of the loading module[80–82]. KS^Q domains are labelled as such because their conserved active-site cysteine is replaced with glutamine[65] (described in more detail in **Section 1.3.3**). As the loading module of DEBS1 lacks a KS^Q domain, it is unable to decarboxylate (2*S*)-methylmalonyl-CoA so must incorporate propionyl-CoA directly.

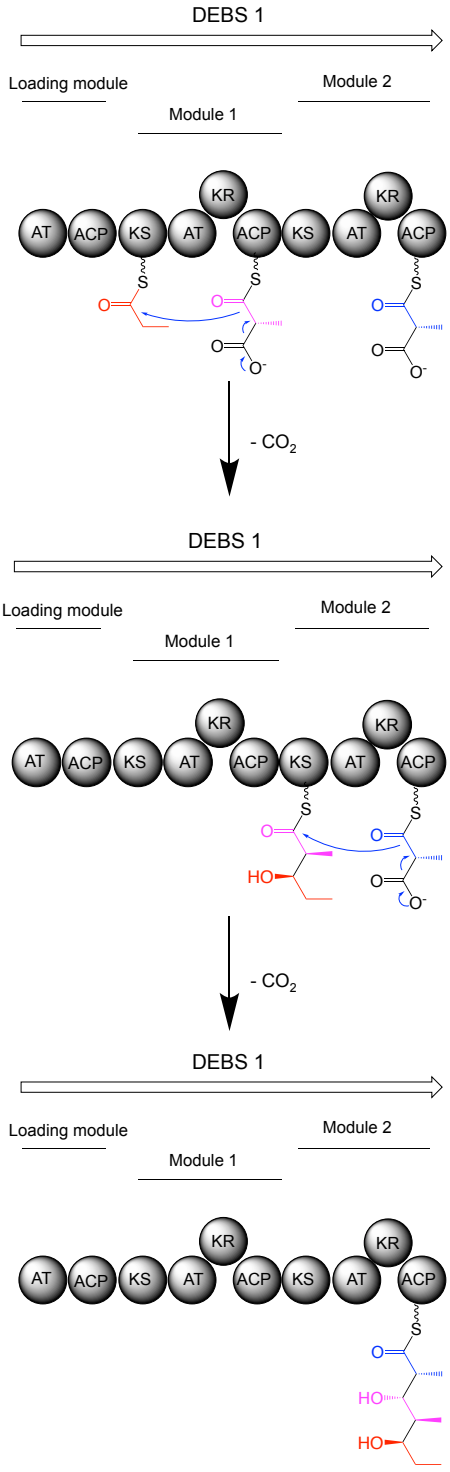


Figure 1.5 The first two extension cycles of DEBS1. DEBS1 is the first type I PKS responsible for the biosynthesis of 6-deoxyerythronolide-B, the erythromycin precursor. The AT domain of the loading module selects a propionyl-CoA unit while the AT domains of modules 1 and 2 select (2*S*)-methylmalonyl-CoA. The KS domains catalyse successive decarboxylative Claisen-like condensation reactions to form a triketide that is subsequently translocated to DEBS2.

The propionate unit bound to the ACP domain of the loading module is translocated to the KS domain of Module 1. The KS domain of module 1 then catalyses the Claisen

condensation reaction between the KS-bound propionate unit and a (2*S*)-methylmalonyl unit bound to the ACP domain of module 1 (**Figure 1.5**). The stereochemistry of the L-α-methyl of the (2*S*)-methylmalonyl-ACP undergoes a stereochemical inversion during the condensation reaction (changing from the L configuration to the D configuration)[83,84] (**Figure 1.6**). Modules 1 and 2 contain a KR domain that catalyses the stereospecific reduction of β-keto groups (either to the D or L configuration). In addition, KR domains may also epimerise the α-substituent.

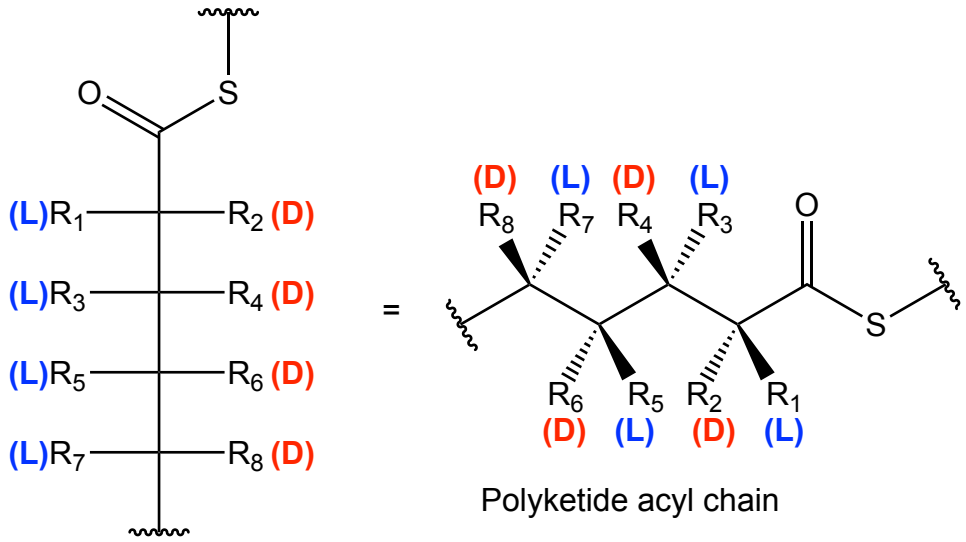


Figure 1.6 D and L configuration of substituents on polyketide chains. It is common to refer to substituents on a polyketide chain as either in the D or L configuration (rather than *R* or *S*). Figure adapted from[65].

For instance, the KR domain of DEBS module 1 produces a D-configured β-hydroxyl and an L-configured α-methyl. The diketide 6-deoxyerythronolide-B intermediate is then translocated to the KS domain of module 2, where it is condensed with another (2S)-methylmalonyl unit bound to the module 2 ACP domain. After processing by module 2, the triketide intermediate of 6-deoxyerythronolide-B intermediate is translocated to the KS domain of module 3, located on DEBS2. A *C*-terminal docking domain on DEBS1 and a *N*-terminal domain on DEBS2 selectively bring these two PKS enzymes together to facilitate this intermolecular translocation[85]. The triketide intermediate is sequentially translocated along extension modules 3-6 of DEBS2 and DEBS3, undergoing an additional chain extension reaction at each, until it reaches extension module 6. Module 6 contains a terminal TE domain that releases the polyketide acyl chain and catalyses a cyclisation reaction to form the 14-membered 6-deoxyerythronolide-B macrolide. Following release from the PKS, other enzymes present in the BGC catalyse further modifications of 6-

deoxyerythronolide-B (**Figure 1.7).** Post-PKS enzymes catalyse two hydroxylations, two glycosylations, and one methyl transfer event to convert 6-deoxyerythronolide-B into erythromycin A[86]. The mechanism of each domain type in type I PKS modules will now be discussed in more detail.

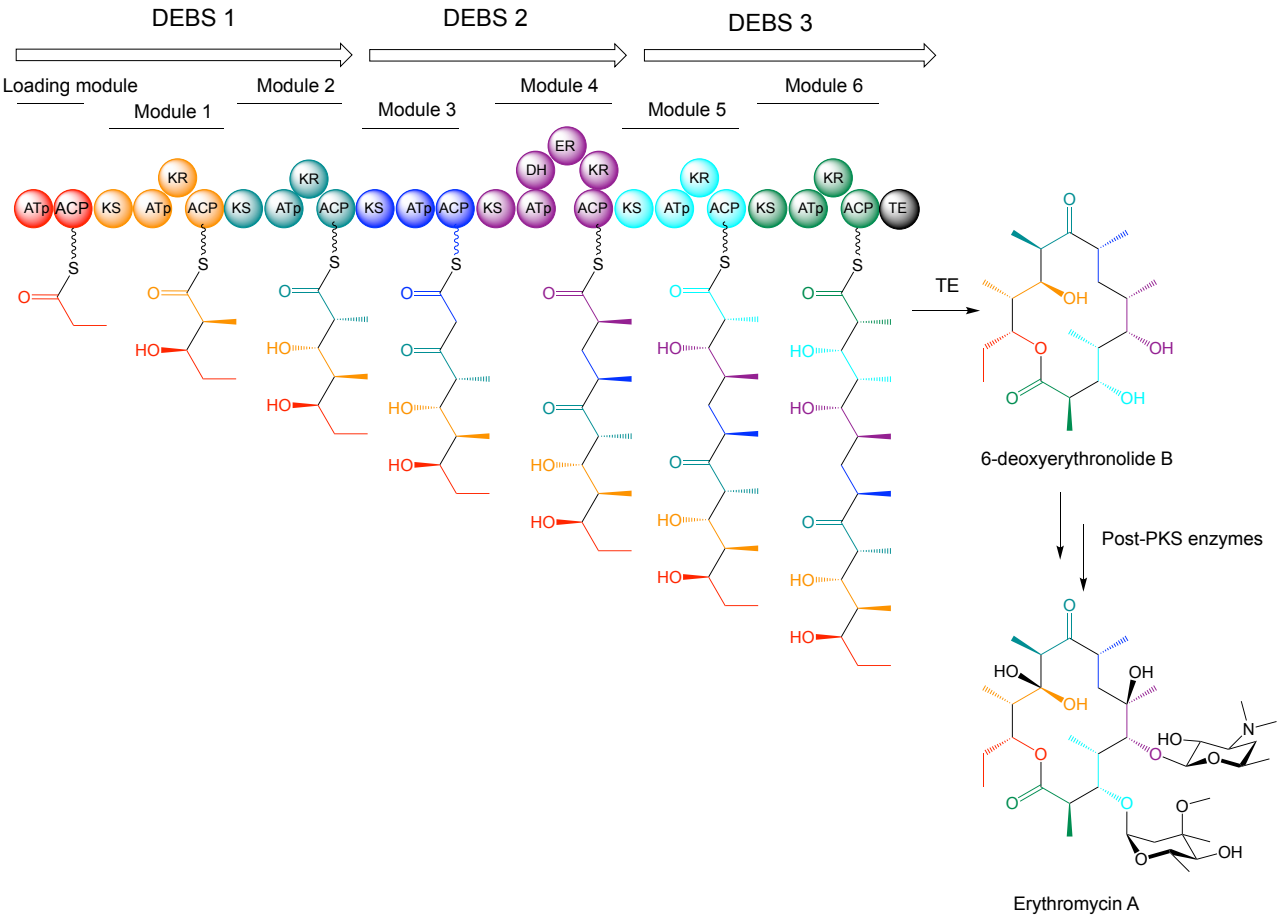


Figure 1.7 Type I PKS biosynthesis pathway of the erythromycin precursor 6-deoxyerythronolide-B. The DEBS complex consists of three type I PKS multienzymes that collectively consist of a loading module and six extension modules. The heptaketide linear 6-deoxyerythronolide-B synthesised by DEBS1-3 is cyclised into a 14 membered macrolide by the terminal TE domain of DEBS3. Post-PKS enzymes then catalyse two hydroxylations, two glycosylations, and one methyl transfer event to produce erythromycin A. Adapted from reference[57].

1.3 Domains of type I polyketide synthases

1.3.1 AT domains

AT domains are responsible for selecting the starter unit and each extender unit used for type I polyketide biosynthesis[65,87]. AT domains in *cis*-PKS pathways (where the AT domains are part of the module) are typically ca. 300 amino acids in size[65]. Most AT domains are highly selective for either malonyl-CoA or (2*S*)-methylmalonyl-CoA, though other extender units such as (2*S*)-ethylmalonyl-CoA or (2*R*)-methoxymalonyl-ACP are also recognised by some AT domains[61]. Numerous substrate motifs within AT domains have been identified that can be used to predict their substrate selectivity (**Table 1.1**). Distinct motifs have also been detected for less common extender units, but they are less well-defined and less predictable[88]. While AT domains selective for malonyl-CoA are able to exclude the bulkier (2*S*)-methylmalonyl-CoA, AT domains selective for (2*S*)-methylmalonyl-CoA can discriminate against the (2R) -epimer but cannot exclude malonyl-CoA[65]. Rather, it has been shown that (2*S*)-methylmalonyl-CoA selective AT domains hydrolyse incorrect extension units faster than the correct one[89]. The active site of AT domains contain a Ser-His catalytic dyad that attacks the thioester carbonyl of the extender unit (**Figure 1.8**). A conserved arginine is then proposed to form a stabilising salt bridge with the α-carboxylate group of the extender unit[90]. The extender unit is then transferred to the adjacent ACP.

AT motif I	**Selectivity**	**Source**
ETGYA(7*x)QxAxFGLL	Malonyl-CoA (acetate)	91
RVDVV(7*x)MxSxAxxW	(2*S*)-methylmalonyl-CoA (propionate)	91
AT motif II		
(HTVY)AFH	Malonyl-CoA (acetate)	92
(YVW)ASH	(2*S*)-methylmalonyl-CoA (propionate)	92

Table 1.1 The different selectivity motifs present in AT domains. Selectivity motifs present in the AT domains of type I PKSs can be used to predict their substrate selectivity. "x" represents any of the proteinogenic amino acids.

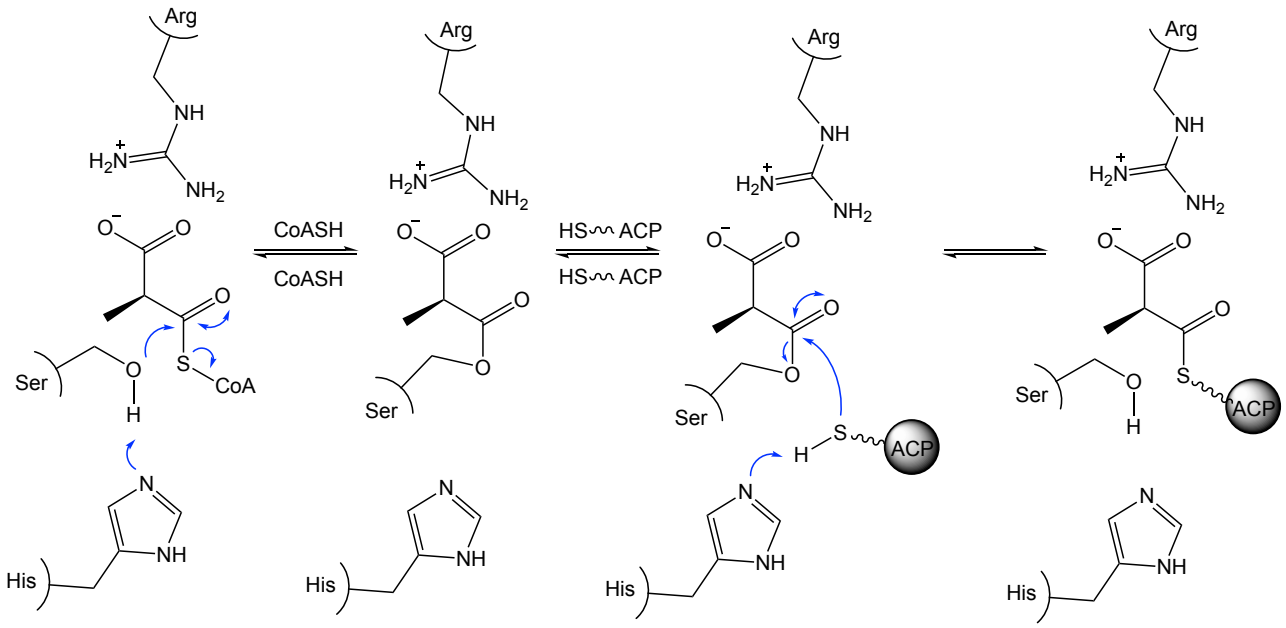


Figure 1.8 The proposed mechanism of AT domains. The active site of an AT domain contains a His-Ser catalytic dyad and a conserved arginine residue. The wavy bond to the ACP represents the 4'-phosphopantetheine group. Figure adapted from[65].

1.3.2 ACP domains

Acyl carrier protein (ACP) domains are typically ca. 80 amino acids in length and found on the *C*-terminal ends of type I PKS modules. ACPs are non-enzymatic domains bearing a covalently attached 4'-phosphopantetheine moiety that tethers the extender unit selected by the AT domain of the same module. Each 4'-phosphopantetheine moiety ("arm") is attached to a conserved serine in the ACP ((D/E)xGxD**S**L), producing *holo*-ACP. 4'-phosphopantetheine is a long molecule (18 Å) derived from co-enzyme A that is attached to the ACP by a 4'-phosphopantetheinyltransferase (PPTase) enzyme[93] (**Figure 1.9**). Due to the necessity of generating *holo*-ACP for polyketide biosynthesis and analogous primary metabolic processes such as fatty acid biosynthesis, PPTases are attractive targets for new antibiotics and efforts have been made to find inhibitors[94–96]. The 4'-phosphopantetheine arm allows the incoming extender unit and the elongated polyketide chain to be tethered as relatively reactive thioesters; and its length and mobility enables it to shuttle polyketide intermediates between the catalytic centers of multiple domains within the current module; as well as transferring the processed chain to the neighboring module for further extension[65,97].

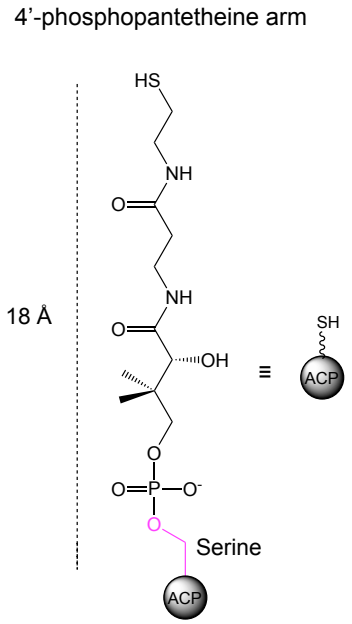


Figure 1.9 The 4'-phosphopantetheine moiety of ACP domains. The 18 Å long 4'-phosphopantetheine moiety is attached to the conserved serine of an ACP domain by a PPTase enzyme. Figure adapted from[65].

1.3.3 KS domains

The KS domain completes the minimal set of type I PKS domains (AT, ACP, and KS) required for polyketide chain extension to occur. KS domains are typically present on the *N*-terminal side of a type I PKS module and are ca. 430 amino acids in size[65]. The main role of the KS is to recruit the polyketide acyl chain from the previous module and catalyse a chain extension reaction between this and a carboxylated extender unit[65]. The active site of KS domains contains a conserved Cys-His-His catalytic triad[98]. The conserved cysteine acts as a nucleophile that accepts the nascent polyketide acyl chain from the ACP domain of the upstream module, while the two histidine residues promote polyketide chain translocation from the ACP to the KS and extender unit decarboxylation[65,99] (**Figure 1.10**). KS-catalysed decarboxylation of the ACP-bound extender unit produces an enolate, the α-substituent of which undergoes inversion of stereochemistry during the subsequent Claisen condensation reaction (L to D transition)[83,84]. KS domains are also known to possess substrate selectivity, only catalysing the condensation if the correct polyketide substrate is bound[100].

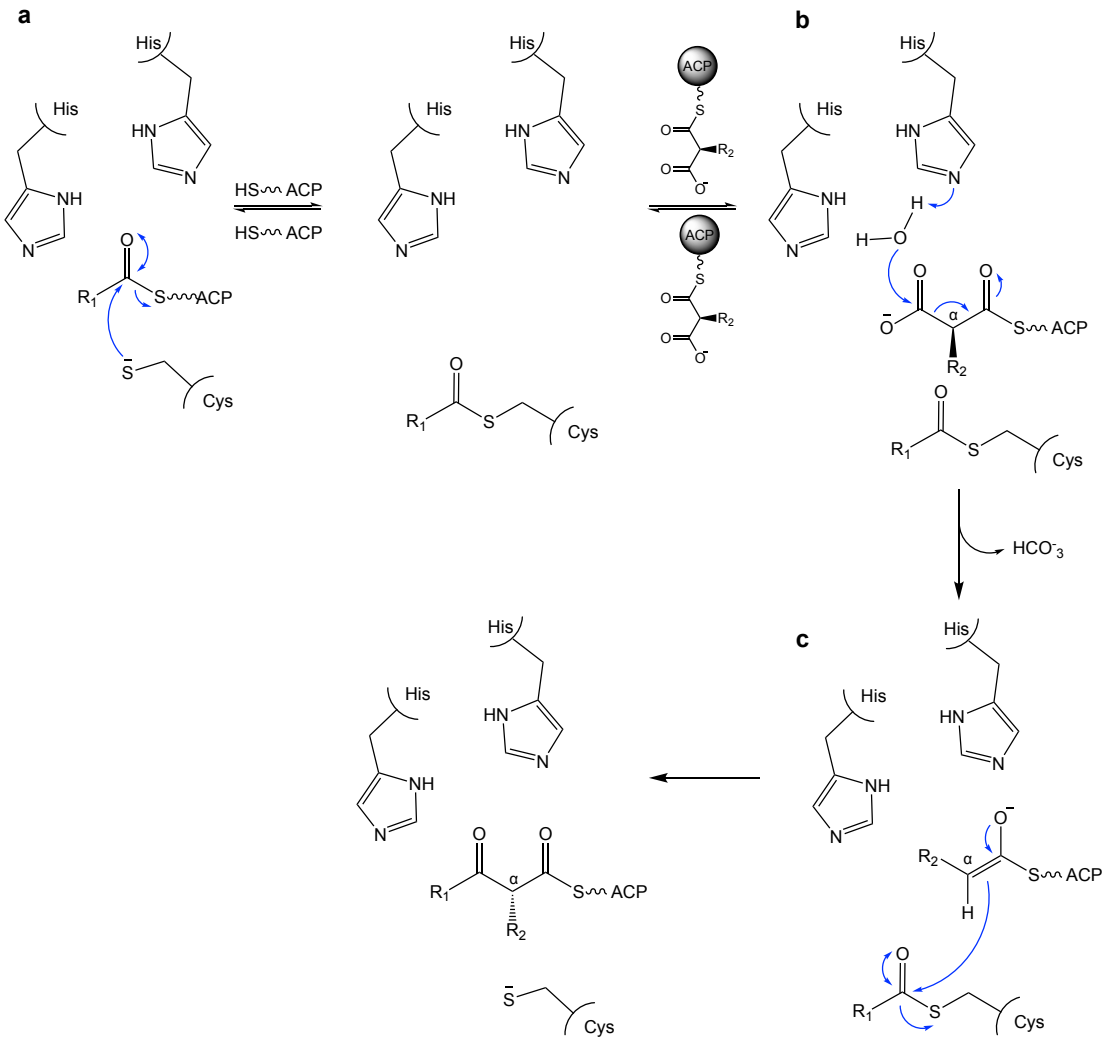


Figure 1.10 The proposed mechanism of KS domains. The active site of KS domains in type I PKSs contain a Cys-His-His catalytic triad. **a,** The polyketide intermediate bound to the upstream ACP domain is translocated to the KS catalytic cysteine. R_1 = polyketide acyl chain **b,** Decarboxylation of the extender unit bound to the downstream ACP produces an enolate. **c,** The enolate then undergoes a condensation reaction with the KS bound polyketide intermediate, extending the polyketide chain by two carbon units The α-substituent of the extender unit undergoes inversion of stereochemistry during the condensation reaction. Figure adapted from[65].

1.3.4 KR domains

KR domains reduce the β-keto group of the ACP-bound polyketide acyl chain to a secondary hydroxyl using an NADPH hydride donor. Conserved serine, lysine, and tyrosine residues are predicted to be essential for catalysis, the conserved tyrosine being the source of the hydroxyl proton (**Figure 1.11**)[65]. Depending on the nature of the KR domain, the secondary hydroxyl may adopt a D or L configuration. Furthermore, KR domains can

also epimerise the α-substituent, making them the domain with the most control over the stereochemistry of the final polyketide product. Specific amino acid motifs within the active site have been used to classify KRs: those that form L-configured hydroxyl groups are denoted A-type, D-configured hydroxyl groups as B-type. KRs lacking epimerase activity are placed in subclass 1, while others that epimerise the α-substituent as well as reducing the β-keto group are placed in subclass 2[101] (**Figure 1.12**). The final class of KR (C-class) is unable to reduce β-keto groups due to mutations in the catalytic tyrosine (C1) or the NADPH binding site (C2), but may retain the ability to isomerise at the α-position (type C2)[101].

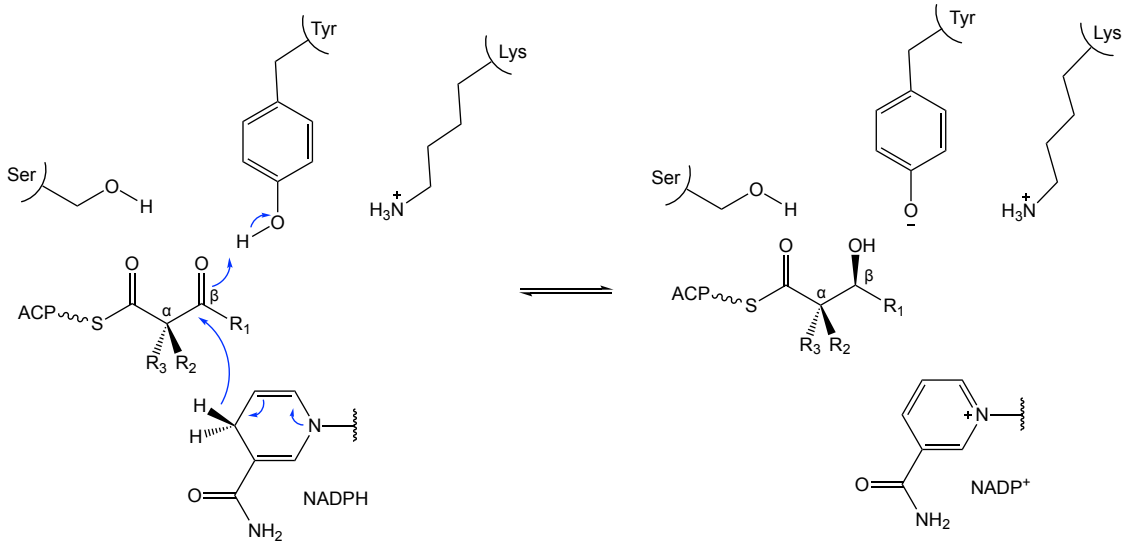


Figure 1.11 The proposed mechanism of KR domains. The active site of a KR domain contains a conserved serine, tyrosine, and lysine all predicted to assist with the reduction of the β-keto group. NADPH acts a hydride donor. In addition, the KR domain may also epimerise the α-substituent to either a D or L configuration. R_1 = polyketide acyl chain. Figure adapted from[65].

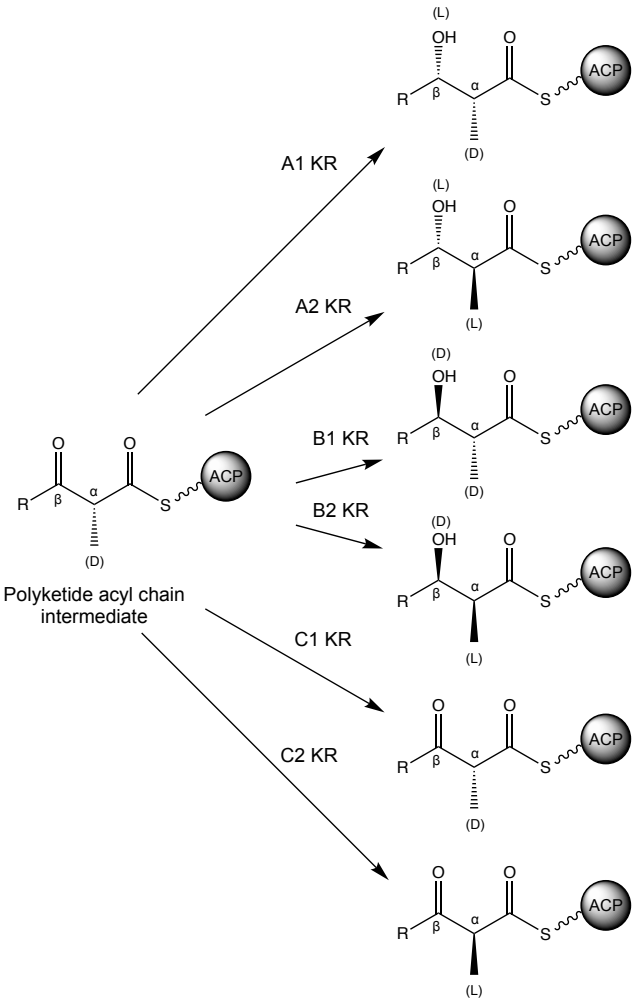


Figure 1.12 The stereochemical outcomes of the different classes of KR domains. There are six different kinds of KR domains, each type producing a different stereochemical outcome at the α and β positions of the polyketide acyl chain. R = polyketide acyl chain. Figure adapted from[101].

1.3.5 DH domains

Following ketoreduction by a KR domain, the ACP-bound polyketide acyl chain is delivered to the active site of a dehydratase (DH) domain if one is present in the module. DH domains are ca. 280 amino acids and catalyse dehydration of the β-hydroxyl generated by the KR domain, forming an α-β double bond[65]. The α-β double bond generated is typically in *trans* configuration, but *cis* double bonds do form in some polyketide biosynthesis pathways. Most double bonds in polyketide natural products are the result of DH domains, although some are produced by post-PKS acting enzymes[102] or, unexpectedly, by some thioesterase (TE) domains[103]. The active site of DH domains contains a His-Asp catalytic dyad. The His residue acts as a general base to abstract the α-proton of the polyketide acyl

chain while the Asp acts as an acid to protonate the β-hydroxyl group, facilitating dehydration by creating a better leaving group[65] (H_2O) (**Figure 1.13**). Whether a DH will form a *cis* (*Z*) or a *trans* (*E*) double bond is not yet possible to predict from the DH sequence alone. However, those that form *cis* double bonds are typically in the same module as an A-type KR, and the evidence suggests that these two modules are collectively responsible for *cis* double bond formation[65,104].

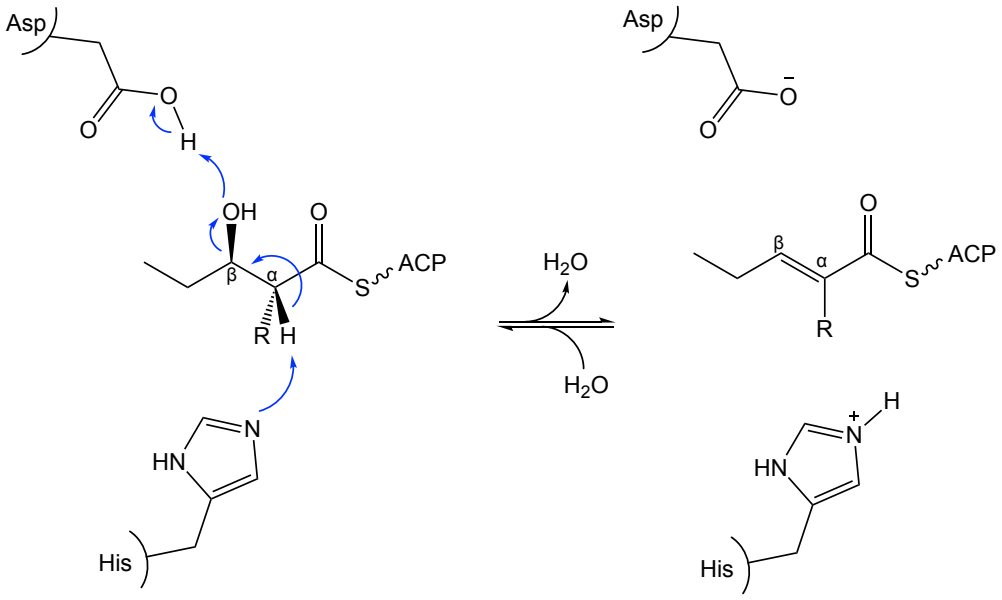


Figure 1.13 The proposed mechanism of the DH domain. The active site of DH domains contains a conserved His-Asp catalytic dyad. The β-hydroxyl group is protonated by the aspartic acid to make it a better leaving group, enabling formation of the (typically *trans*) α-β double bond. Figure adapted from[65].

1.3.6 ER domains

The final reduction domain optionally present in a type I PKS module is the enoylreductase (ER) domain. ER domains are ca. 310 amino acids in length and reduce the DH-catalysed double bond to a fully saturated carbon-carbon bond[65]. Like the KR domain, ER domain reduction is achieved using NADPH as a cofactor. Hydride donation by NAPDH results in the formation of an enolate (**Figure 1.14**). A proton is then added to either the *re* face or the *si* face of the enolate to produce a product with an L-α or D-α configured substituent, respectively.

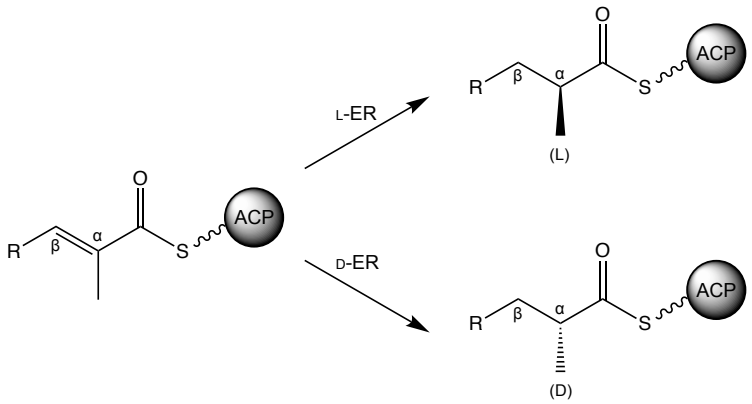


Figure 1.14 The two different classes of ER domain. ER domains reduce the α-β double bond of a polyketide acyl intermediate to a saturated carbon bond. L-ER domains result in the α-substituent being in the L-configuration, whereas D-ER domains result in the α-substituent being in the D-configuration. R = polyketide acyl chain. Figure adapted from[65].

ER domains that produce D-α substituents (D-ER domains) can be distinguished from those that produce L-α substituents (L-ER domains) by specific amino acid motifs[105]. L-ER domains contain a key tyrosine residue that donates a proton to the *re* face of the enolate to produce a L-α substituent (**Figure 1.15**). On the other hand, D-ER domains lack this tyrosine residue and the proton is added to the *si* face, forming a D-α substituent. The source of the proton added to the *si* face is possibly a conserved lysine, but there is also evidence that this proton has multiple sources[105,106].

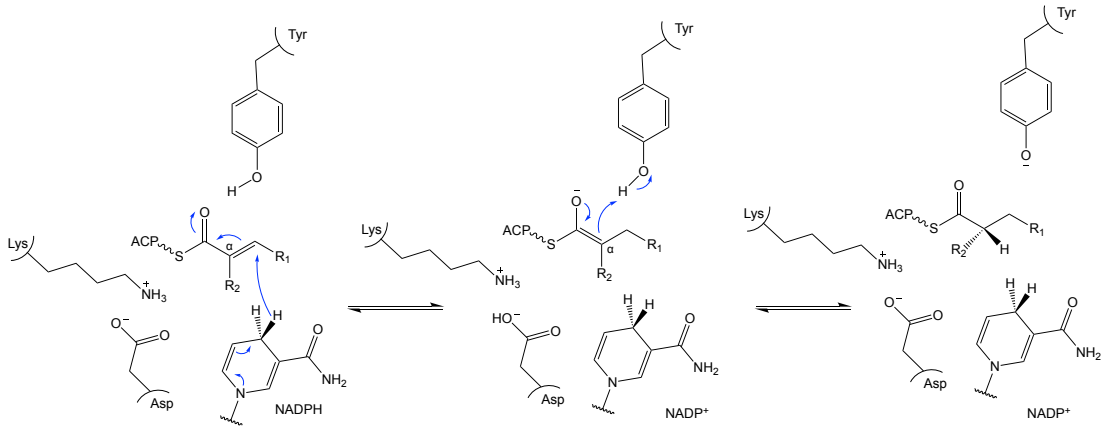


Figure 1.15 Proposed mechanism of the ER domain. ER domains reduce the α-β double bond of the polyketide acyl chain using NADPH as a hydride donor. The figure displays an L-type ER that uses a conserved tyrosine to produce an L-configured α-substituent. R_1 = polyketide acyl chain. R_2 = α substituent (typically a methyl group). Figure adapted from[65].

1.3.7 TE domains

The final domain in most type I PKSs is the thioesterase (TE) domain. TE domains are 240-290 amino acids in size and are typically located at the end of the final extension module (module 6 in the case of the DEBS proteins). The role of the TE domain is to catalyse chain release from the PKS, either as a macrocycle or as a free acid. The active site of a TE domain contains a Ser-His-Asp catalytic triad, with the serine acting as a nucleophile to attack the thioester carbonyl of the ACP-bound polyketide acyl chain (**Figure 1.16**)[65]. While thioesterase domains offer the most common mechanism of chain release, several other mechanisms exist[107,108]. For instance, chain release of fumonisin mycotoxin polyketides is performed by a standalone pyridoxal 5'-phosphate (PLP) dependent enzyme. Another important exception relevant to this thesis are the polyketide tetronate natural products, the PKS enzymes of which lack thioesterase domains. Chain release is instead concomitant with tetronate ring formation[109,110–118].

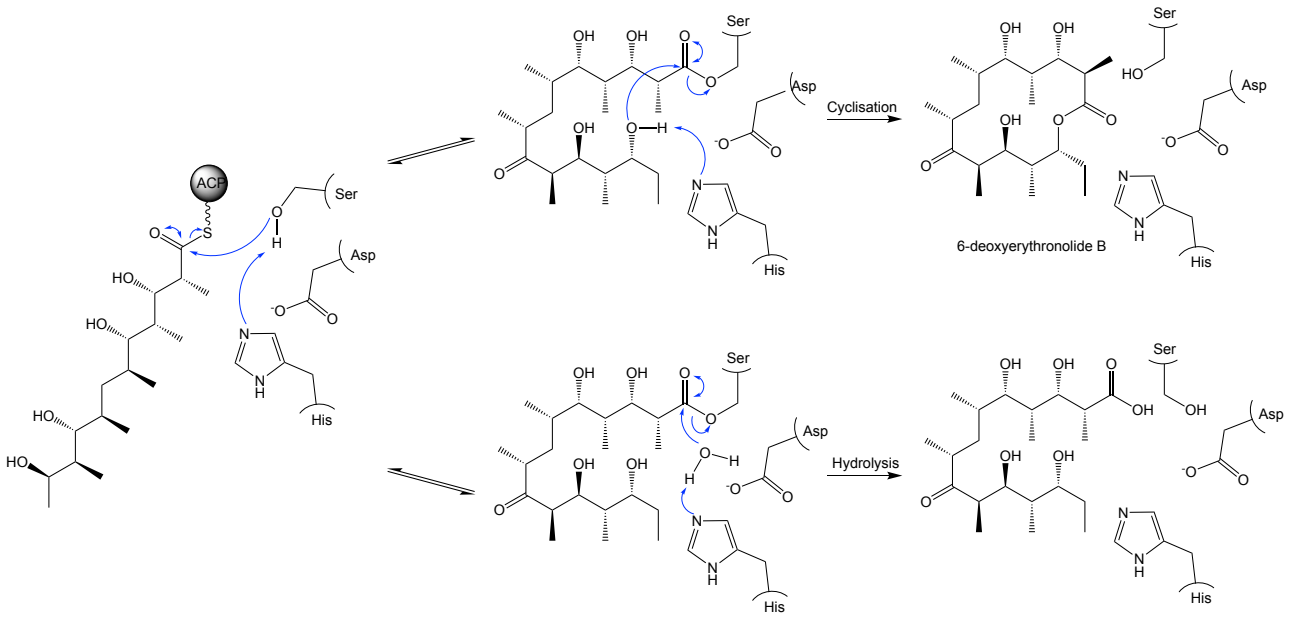


Figure 1.16 Proposed mechanism of TE domains. The active site of a TE domain contains a conserved Ser-His-Asp catalytic triad responsible for catalysing release of the polyketide acyl chain (the figure shows the linear ACP-bound intermediate of 6-deoxyerythronolide-B) either by hydrolysis of intramolecular cyclisation (as is done by the TE domain of DEBS3). Figure adapted from[65]

1.4 Tetronate natural products

Polyketides or fatty acids containing a tetronic acid (**Figure 1.17**) moiety (collectively called tetronates) are a diverse family of natural products. (**Figure 1.18**). The first tetronate natural product was discovered in 1935[119], and since then over 100 have been discovered[109]. Tetronate natural products exhibit a range of activities, including antibiotic, anticancer, antiviral, and antifungal[109]. Most tetronate natural products can be categorised into two broad classes: the linear tetronates and the spirotetronates. Linear tetronates contain a fatty acid or a polyketide chain attached to a tetronate group (examples including sarcotin A, RK-682, and pesthetoxin). The tetronate moiety is a steric and electronic mimic of a phosphate group, consistent with many linear tetronates having phosphatase inhibitor activity[109]. The spirotetronates are macrocyclic polyketides where the tetronic acid moiety is in a spiro linkage to a cyclohexene ring. The first spirotetronate discovered was chlorothricin produced by *Streptomyces antibioticus* Tü99[120]. The size of macrocyclic ring can vary between spirotetronates, ranging from 11-carbons in the abyssomicins to 13+ in the quartromicins and versipelostatin A[109].

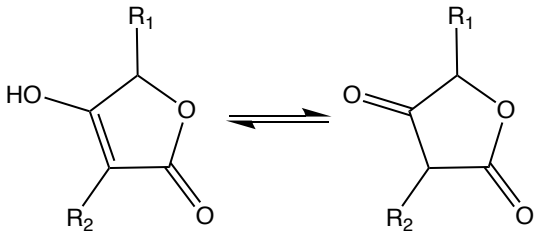


Figure 1.17 Tautomers of tetronic acid.

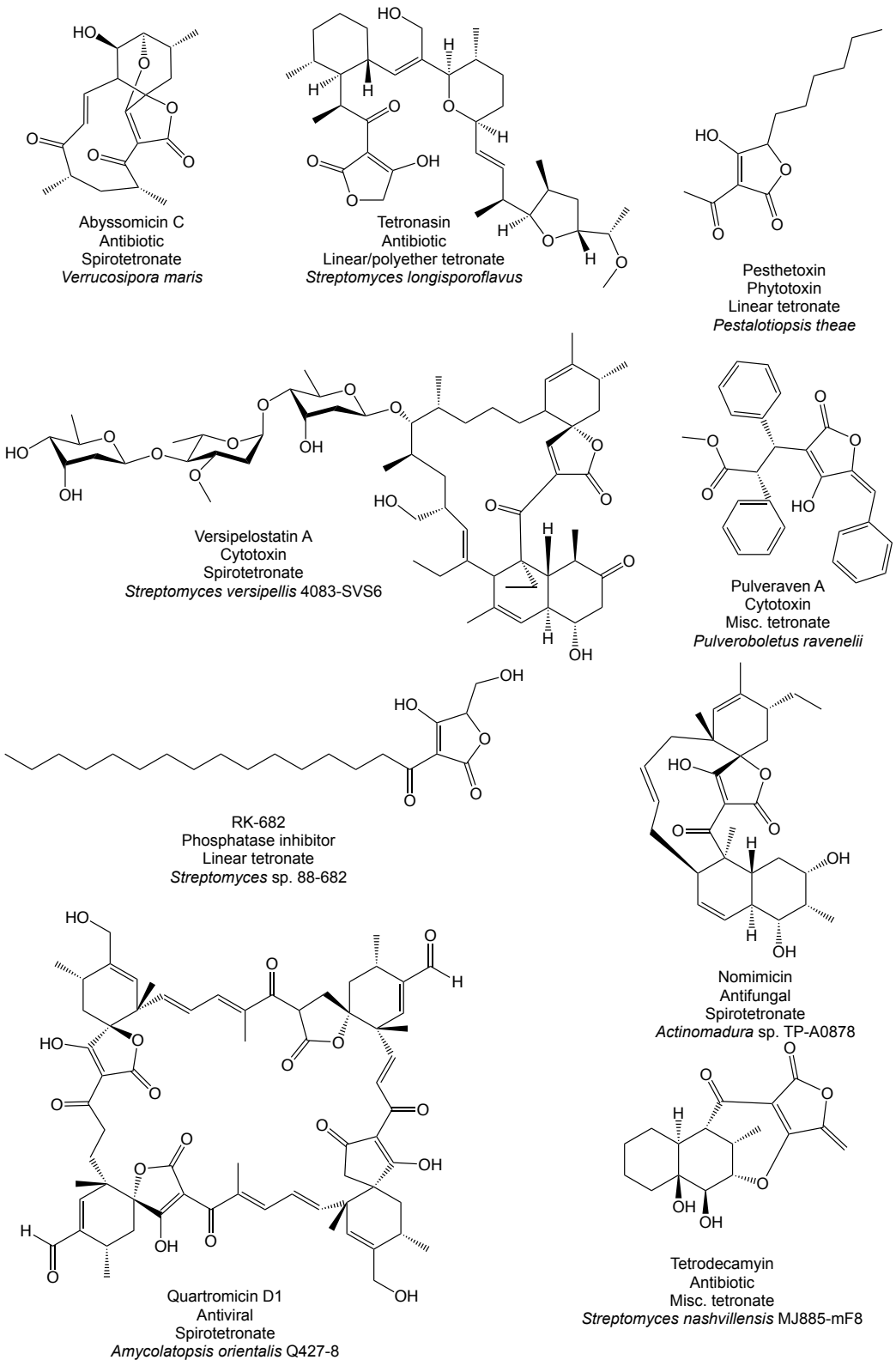


Figure 1.18 Structures of representative polyketide tetronate natural products. Polyketide tetronate natural products are a diverse family of natural products with a range of different activities.

The tetronate moiety itself is normally formed by a reaction between the full length ACP-bound polyketide acyl chain and glyceryl-ACP (derived from 1,3-bisphosphoglycerate)[121]. The mechanism of tetronate formation was elucidated by reconstituting its biosynthesis *in vitro* for the linear tetronate RK-682[122]. The 3-5 conserved genes typically required for tetronate biosynthesis are labelled the "glycerate utilisation operon". The three essential genes in the glycerate utilisation operon are a FkbH-like enzyme, a FabH like enzyme, and a standalone acyl carrier protein[109]. The FkbH enzyme catalyses the formation of glyceryl-ACP from 1,3-bisphosphoglycerate and the standalone acyl-carrier protein[121]. The FabH-like enzyme (homologous to enzymes in fatty acid synthesis that catalyse chain elongation) then catalyses both C-C and C-O bond formation between glyceryl-S-ACP and the polyketide acyl chain, forming the five-membered tetronate ring and releasing the polyketide chain from the PKS[122] (**Figure 1.19**). In polyketide tetronate pathways the full-length ACP-bound polyketide acyl chain always contains a β-keto group to increase the acidity of the α-proton, subsequently increasing the nucleophilicity of the α-carbon to promote C-C bond formation.

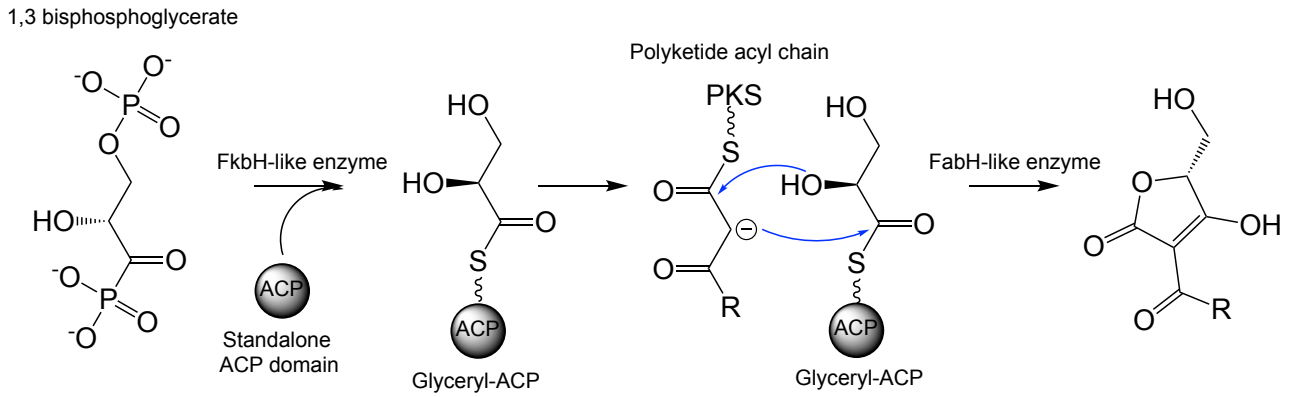


Figure 1.19 The mechanism of tetronate ring biosynthesis. An FkbH-like enzyme catalyses formation of glyceryl-ACP from 1,3 bisphosphoglycerate and a standalone ACP. A FabH-like enzyme then catalyses C-C and C-O bond formation to create the tetronate ring. R = polyketide acyl chain.

The glycerate utilisation operon may contain two additional genes to introduce an exocyclic double bond on the tetronate. In a mechanism elucidated by studying the linear tetronate agglomerin A (Agg), an acyltransferase enzyme (Agg4) was found to catalyse acetylation of the C-6 primary hydroxyl group, followed by elimination of the acetyl group by a

dehydratase-like enzyme (Agg5), forming the exocyclic double bond[123] (**Figure 1.20**). Accordingly, homologues of Agg4 and Agg5 are encoded within the BGCs of all tetronate products that contain (or contained at one point during their biosynthesis) an exocyclic double bond on the tetronate moiety.

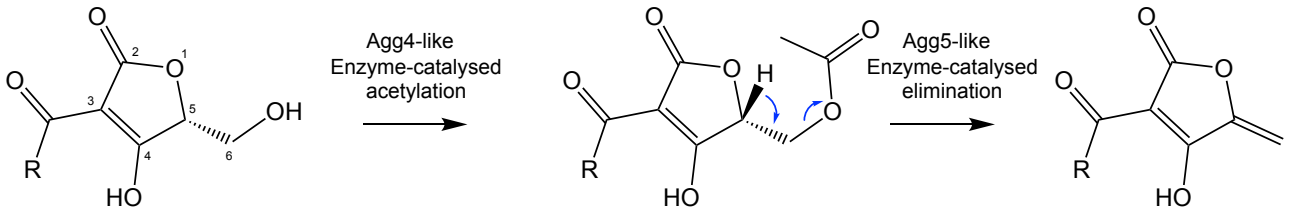


Figure 1.20 Formation mechanism of the tetronate exocyclic bond. An acyltransferase enzyme catalyses acetylation of the C6 primary hydroxyl group. The acetyl group is then eliminated by another enzyme that resembles a dehydratase, forming the exocyclic double bond.

These exocyclic double bonds attracted considerable interest as they were hypothesised to be involved in spirotetronate formation via an intramolecular Diels-Alder reaction ([4+2] cycloaddition) with the terminal 1,3 diene of the polyketide intermediate (**Figure 1.21**)[109].

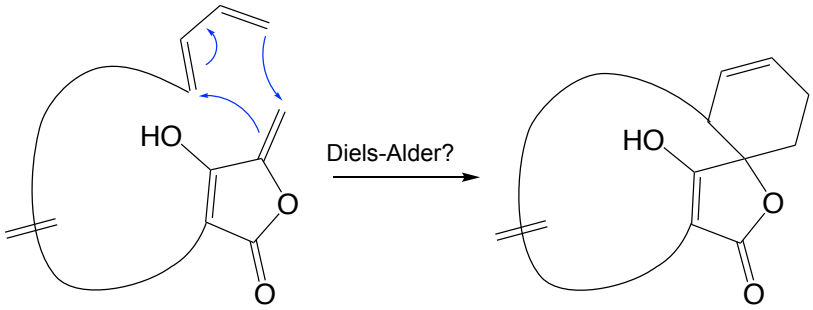


Figure 1.21 Proposed mechanism of spirotetronate formation via an intramolecular Diels-Alder reaction. The exocyclic double bond on the tetronate reacts in an intramolecular Diels-Alder with the 1,3 diene located on the terminal end of the polyketide intermediate.

In its simplest form a Diels-Alder reaction is the reaction between a 1,3 diene and an alkene to form a cyclohexene ring. The Diels-Alder reaction is one of the most important reactions ever discovered for synthetic organic chemistry, so the prospect that an enzyme

might have evolved to catalyse one in a natural product biosynthesis pathway was very exciting. Accordingly, whether biosynthetic Diels-Alder reactions occur and whether enzymes catalyse them has been the subject of much investigation.

1.5 The Diels-Alder reaction

The Diels-Alder reaction was discovered in 1928 by the German chemists Otto Diels and Kurt Alder[124]. The formal description of the Diels-Alder reaction is a [4+2] cycloaddition, where a 1,3 diene and an alkene (also called a dienophile) react via a cyclic transition state to form a cyclohexene ring (**Figure 1.22**). If a heteroatom such as oxygen or nitrogen is incorporated into the six-membered ring product, the reaction is additionally referred to as a hetero-Diels-Alder reaction[125].

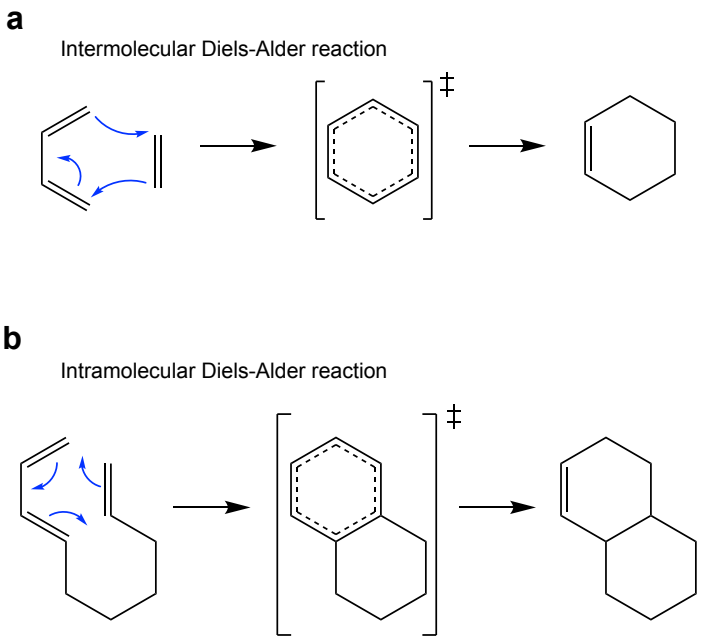


Figure 1.22 Intermolecular and intramolecular Diels-Alder reaction. a, In an intermolecular Diels-Alder reaction two compounds, one a 1,3 diene and the other an alkene, react to form a cyclohexene ring. **b,** In an intramolecular Diels-Alder a single compound that contains both a 1,3 diene and an alkene functionality react to form a cyclohexene ring.

The discovery of the [4+2] cycloaddition earned Diels and Alder the 1950 Nobel prize in Chemistry[126]. A cycloaddition reaction is a member of the pericyclic reaction family, the other major family members being the group transfers, electrocyclic and sigmatrophic

rearrangement reactions. A reaction is classified as pericyclic if it has a cyclic transition state and bond formation and bond breakage occur in concert[125]. The reaction is classified as a [4+2] cycloaddition in reference to the number of π-electrons of the 1,3 diene (4π) and the alkene (2π). The Diels-Alder reaction, and pericyclic reactions in general, are controlled by molecular orbital symmetry of the reagents. In a typical Diels-Alder reaction the highest occupied molecular orbital (HOMO) of the diene reacts with the lowest unoccupied molecular orbital (LUMO) of the dienophile. The phases of the diene HOMO and dienophile LUMO molecular orbitals are symmetrical, such that constructive orbital overlap occurs and new bonds can form (**Figure 1.23a**). By making the relative energy levels of the HOMO and LUMO more similar the rate of Diels-Alder reaction can be increased[125]. Such energy level balancing is typically achieved by attaching an electron-withdrawing group (EWG) to the dienophile (thereby lowering the LUMO energy) and/or adding an electron-donating group (EDG) to the diene (thereby increasing the HOMO energy level)[125] (**Figure 1.23a and b**).

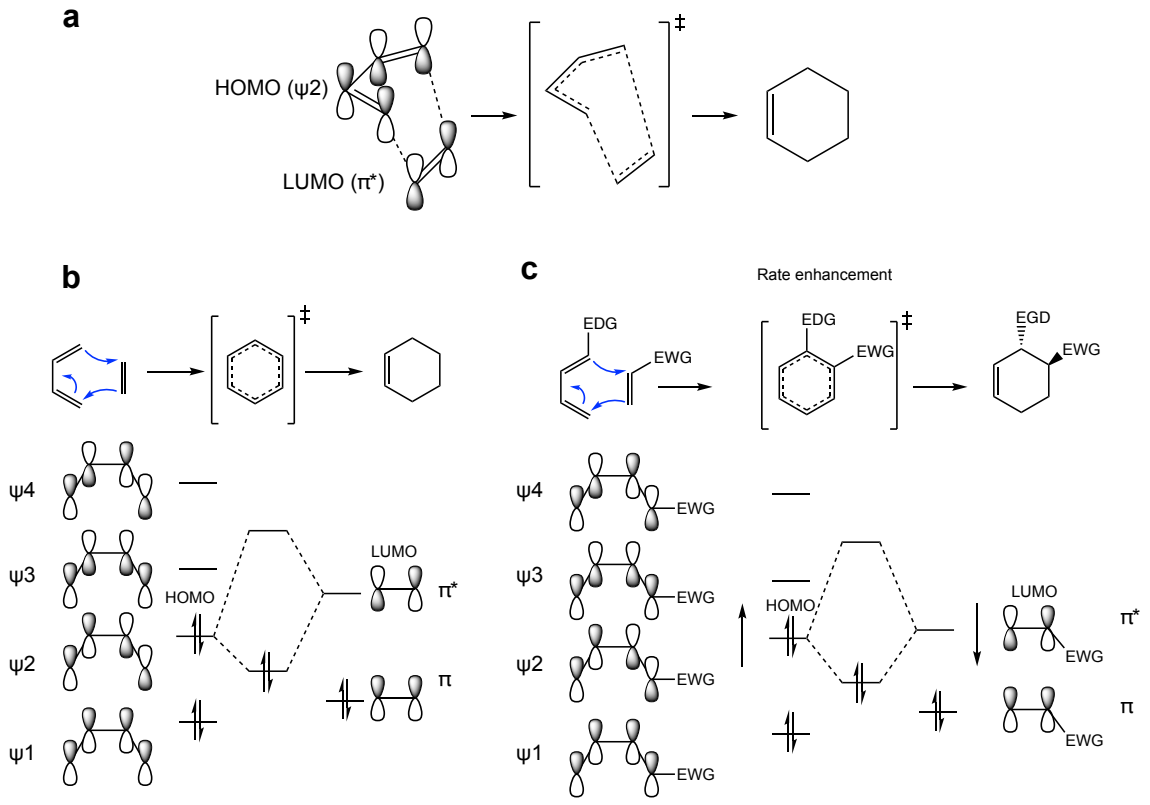


Figure 1.23 Molecular orbital overlap of the Diels-Alder reaction. a, In a typical Diels-Alder reaction the ψ2 molecular orbital (HOMO) of the 1,3 diene reacts with the π* (LUMO) molecular orbital of the alkene. The symmetry of these orbitals enables constructive orbital overlap to occur, resulting in two C-C bonds forming in a concerted step. **b,** Energy level diagram of the molecular orbitals used in the reaction between butadiene and ethene (the simplest Diels-Alder reaction) to form cyclohexene. **c,** Adding electron donating groups (EDG) substituents to butadiene and EWG substituents to the ethene makes the relative energy level of the HOMO and LUMO more similar, enhancing the rate of reaction.

The orientation of the dienophile relative to the diene during the reaction can result in multiple products being produced. The major product of a [4+2] cycloaddition reaction is typically the *endo* product: the product that forms when the EWG substituents of the dienophile are oriented towards the diene (**Figure 1.24**). In contrast, the *exo* product, where the EWG substituents of the dienophile are oriented away from the diene, is typically the minor product. Why the *endo* product is favoured is unclear, but may have to do with stabilising secondary orbital interactions made between the substituents of the dienophile and the diene, though this explanation has been recently challenged[127]. Regardless of why it occurs, the preference for the *endo* product gives the Diels-Alder reaction its true power: the ability to predict the stereochemistry of the four stereocenters in the cyclohexene product.

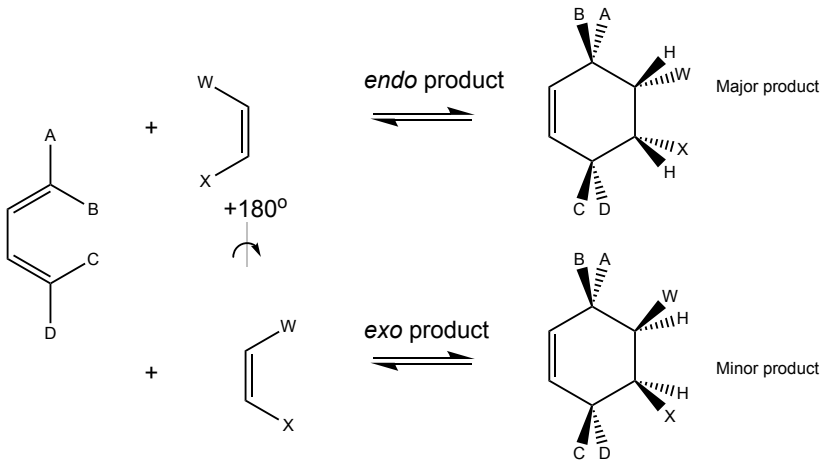


Figure 1.24 *endo* and *exo* products of a Diels-Alder reaction. Depending on the orientation of the dienophile to the 1,3 diene two different products are possible: called the *endo* and an *exo*. Under typical conditions the *endo* product is the major product (an observation known as Alder's *endo* rule).

It is also possible for a Diels-Alder reaction to occur between the HOMO of the dienophile and the LUMO of the diene. Such a reaction is called an inverse-electron-demand Diels-Alder reaction and is only possible if there is an EWG on the diene (such as an oxygen atom) and/or an EDG on the dienophile (**Figure 1.25**). While the Diels-Alder reaction is a staple in synthetic organic chemistry, whether [4+2] cyclases have evolved has been the subject of much investigation.

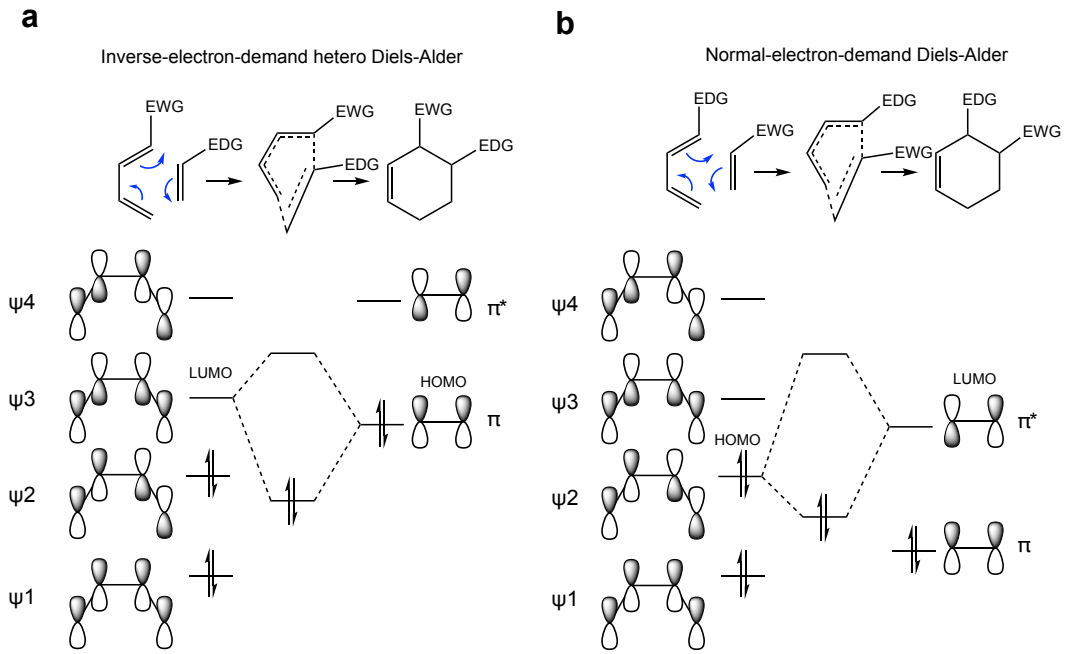


Figure 1.25 The molecular orbitals used in an inverse-electron-demand Diels-Alder reaction. a, In an inverse-electron-demand Diels-Alder reaction the LUMO of the 1,3 diene (ψ3) reacts with the HOMO of the dienophile (π). Adding an electron-withdrawing group (EWG) to the diene and/or an electron-donating group (EDG) to the diene can enhance the reaction. **b,** A normal demand Diels-Alder reaction for comparison.

1.6 The hunt for [4+2] cyclase enzymes

Initial efforts to find a [4+2] cyclase focused on creating one, rather than sourcing one from nature. The first synthetic [4+2] cyclases were monoclonal antibodies. To achieve this, monoclonal antibodies were raised to bind a small molecule that mimicked the cyclic transition state of a specific [4+2] cycloaddition reaction. The rationale was that the antibody would bring the diene and dienophile together in a reactive conformation where the molecular orbital overlap required for the [4+2] cycloaddition could occur. Furthermore, by raising the antibody to a transition state mimic, rather than the product, product inhibition could be avoided[128]. Using this method catalytic antibodies were created that could enhance the rate and product selectivity (*endo* vs *exo*) of a diverse range of [4+2] cycloaddition reactions[128–132]. In addition to catalytic antibodies, a ribozyme (RNA) [4+2] cyclase was also discovered by screening a $2x10^{14}$ library of ribose nucleotides[133]. Despite the flexibility of these approaches, they have fallen out of favour due to the rate enhancements they provide being poor compared to natural enzymes[134]. In an attempt to design a superior “non-natural” [4+2] cyclase, *de novo* enzyme design was successfully

applied to create a synthetic [4+2] cyclase[135]. Analysis of the active site of the *de novo* [4+2] cyclase revealed it was largely a hydrophobic pocket with containing a glutamine (EDG) and a tyrosine (EWG) that were proposed to respectively increase the HOMO energy of the diene and decrease the LUMO energy of the dienophile[135].

There are now numerous examples of natural products that apparently require a [4+2] cycloaddition reaction in their biosynthesis, and in several cases putative [4+2] cyclase enzymes have been identified[136,137]. Whether these enzymes catalyse true [4+2] cycloaddition reactions has been subject to much debate, since it is difficult to prove whether the reactions are truly concerted. Reference to one of these enzymes as a [4+2] cyclase or as a Diels-Alderase is therefore done largely out of convenience to describe an enzyme that catalyses a reaction between a 1,3 diene and an alkene to form a cyclohexane ring[137]. Furthermore, the initial [4+2] cyclase candidates all also possess an additional, non-cyclase, activity and/or they catalyse [4+2] cycloadditions that also occurs spontaneously[138].

1.6.1 Dual function [4+2] cyclases

The first natural product biosynthetic pathway to provide direct evidence of a [4+2] cycloaddition reaction occurring was that of the solanopyrones[139] (**Figure 1.26**). Solanopyrone A and D are phytotoxic polyketides produced by the *Alternaria solani* fungus[140]. From the time of their discovery the dialkyldecalin moiety of the solanopyrones was predicted to form via an intramolecular [4+2] cycloaddition reaction, solanapyrone A being the *exo* product and solanapyrone D being the *endo* product. The key experiment that demonstrated a Diels-Alder reaction involved feeding *Alternaria solani* isotopically labeled prosolanapyrone II, the proposed precyclisation precursor of the solanopyrones (**Figure 1.26**) . The isotopically labelled prosolanopyrone II was incorporated into solanopyrone A, strongly suggesting that it had undergone the predicted intramolecular Diels-Alder reaction[139]. A cell free extract from *Alternaria solani* was found capable of oxidising the primary hydroxyl of prosolanapyrone II to an aldehyde, forming prosolanapyrone III[141]. Prosolanapyrone III spontaneously undergoes an apparent [4+2] cyclisation to form the *exo* and *endo* products at a 3:97 ratio, respectively[141]. However, the cell-free extract, later shown to contain the enzyme Sol5[142], accelerated the cyclisation

nearly four fold and selectively produced the *exo* product over the *endo* at a 7:1 ratio[141,142]. Notably, when purified Sol5 degrades over time, its [4+2] cyclase activity decreases faster than the oxidase activity, suggesting specific protein elements are responsible for each[143]. Unfortunately no Sol5 crystal structure is available to investigate this proposal further. Whether Sol5 is a true [4+2] cyclase is still subject to debate namely because it has monooxygenase activity and the reaction it catalyses also occurs spontaneously.

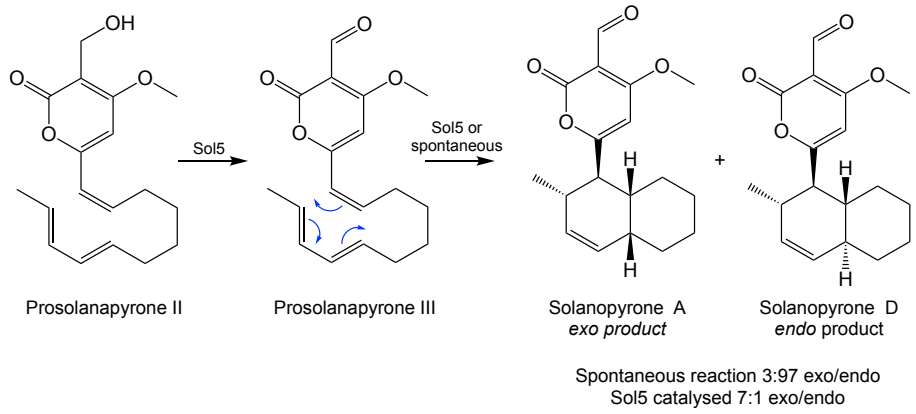


Figure 1.26 Solanopyrone biosynthesis. The enzyme Sol5 oxidises the primary hydroxyl of prosolanapyrone I to an aldehyde, forming prosolanapyrone III. Prosolanapyrone III can spontaneously undergo an apparent [4+2] cycloaddition to form the *exo* and *endo* products at a 3:97 ratio. In the presence of Sol5 the exo/endo ratio changes to 7:1 and the rate of the reaction is enhanced four-fold. Figure adapted from[136]

Another notable early example of the difficulties in confidently ascribing [4+2] cyclase activity to an enzyme is the LovB/LovC complex from the lovastatin biosynthesis pathway. Lovastatin is a fungal polyketide natural product used for the treatment of hyperlipidemia and the prevention of cardiovascular disease[144]. Like solanapyrone A/D, lovastatin also contains a dialykyldecalin predicted to form via a [4+2] cycloaddition (**Figure 1.26**). The lovastatin BGC was first sequenced from *Aspergillus terreus* and found to consist of two iteratively acting type I PKS enzymes, LovB and LovF. Together these PKS enzymes are responsible for synthesising the nonaketide and diketide portions of lovastatin, respectively[145] (**Figure 1.27**). LovB forms a complex with LovC, an enoyl-reductase protein required for biosynthesis of the nonaketide[146]. Intriguingly, the nonaketide offloaded from LovB already contains the dialkyldecalin moiety, suggesting that an intramolecular [4+2] cycloaddition reaction had taken place during the iterative LovB-catalysed polyketide

extension cycles[146,147]. A chemically synthesised *N*-acetylcysteamine (a short 4'-phosphopantithine mimic) hexaketide analogue of the predicted pre-cyclisation lovastatin intermediate was shown to spontaneously undergo an apparent [4+2] cycloaddition reaction to form a mix of *endo* and *exo* products. However, none of the products had the same ring stereochemistry as lovastatin[137,148]. On the other hand, when this substrate was incubated with purified LovB, the correct dialkyldecalin stereochemistry formed, albeit only as a minor product[137,148]. How LovB functions *in vivo* to selectively produce the correct decalin ring stereochemistry of lovastatin remains unknown, but it is another intriguing example of a putative duel function [4+2] cyclase enzyme.

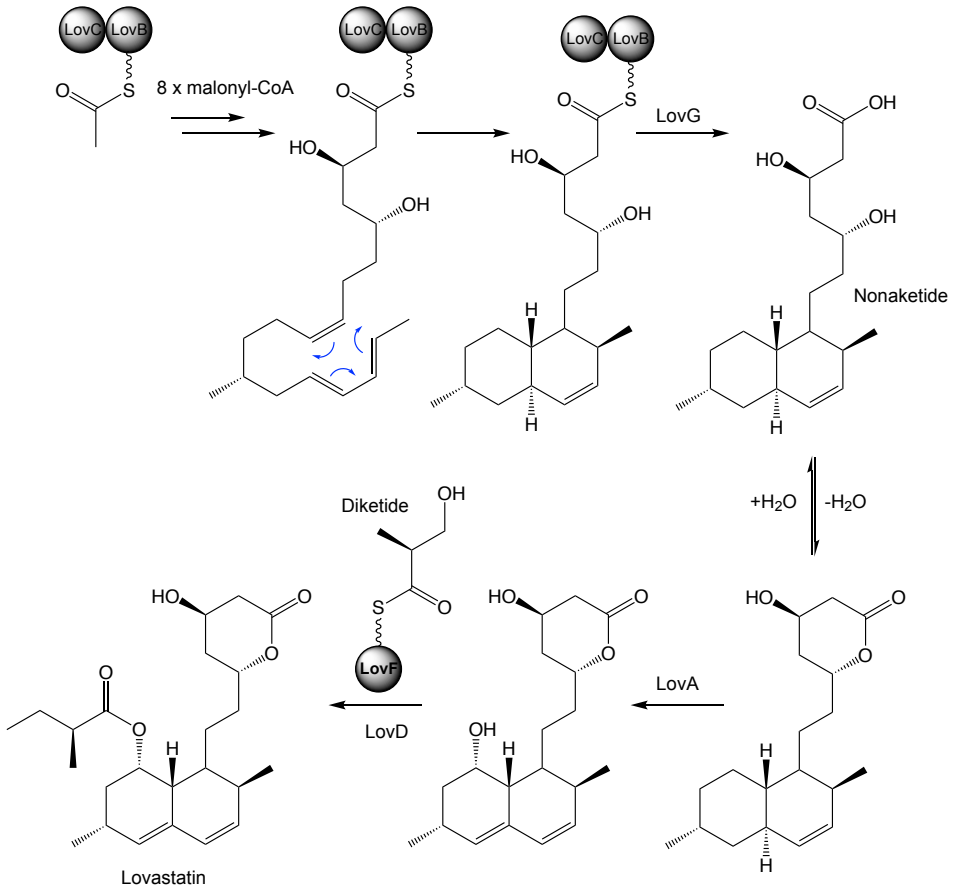
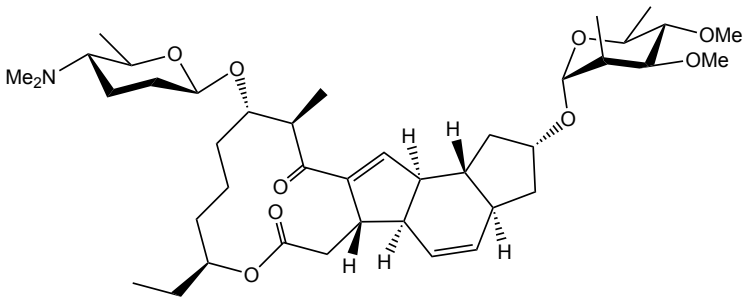


Figure 1.27 Lovastatin biosynthesis. The iterative type I PKS LovB synthesises the linear nonaketide lovastatin precursor. The nonaketide intermediate undergoes an apparent intramolecular [4+2] cycloaddition while still attached to the LovB/LovC complex, suggesting this enzyme complex could have [4+2] cyclase activity. Figure adapted from[137]

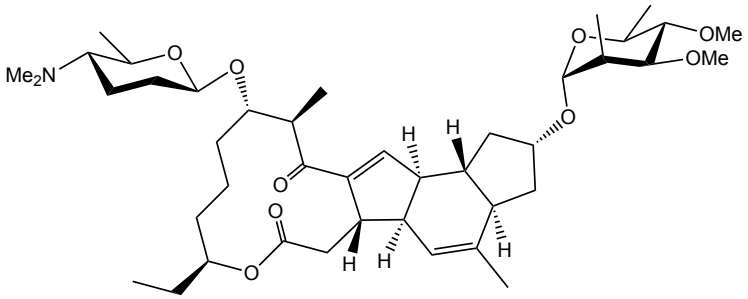
1.7 Monofunctional [4+2] cyclases

1.7.1 SpnF

SpnF from the spinosyn biosynthesis pathway was the first documented example of an enzyme whose single function appeared to be a [4+2] cycloaddition[149]. Spinosyn A and D (**Figure 1.28**) are type I polyketide insecticides produced by *Saccharopolyspora spinosa*[150,151], an actinomycete isolated from the soil beneath a rum distillery in the British Virgin Islands[152].



Spinosyn A



Spinosyn D

Figure 1.28 Structure of spinosyn A and D. Spinosyn A and D are polyketide natural products produced by *Saccharopolyspora spinosa*.

The spinosyns consist of a 12 membered macrolactone connected to a perhydro-*as*-indacene moiety. The latter was predicted to form, in part, via a [4+2] cycloaddition reaction[153]. In contrast to the multifunctional Sol5 or LovB/LovC, a standalone S-adenosylmethionine-binding enzyme, SpnF, appeared to catalyse this [4+2] cycloaddition reaction (although alternative cyclisation mechanisms cannot be entirely ruled out)[149,154] (**Figure 1.29**). While this apparent [4+2] cyclisation process can also occur spontaneously, SpnF enhances its rate 500 fold[149], making it the best candidate yet for a *bona fide* [4+2] cyclase. The crystal structure of SpnF was subsequently solved to 1.5 Å, though its

mechanism remains largely mysterious (discussed in more detail in **Chapter 6**)[155].

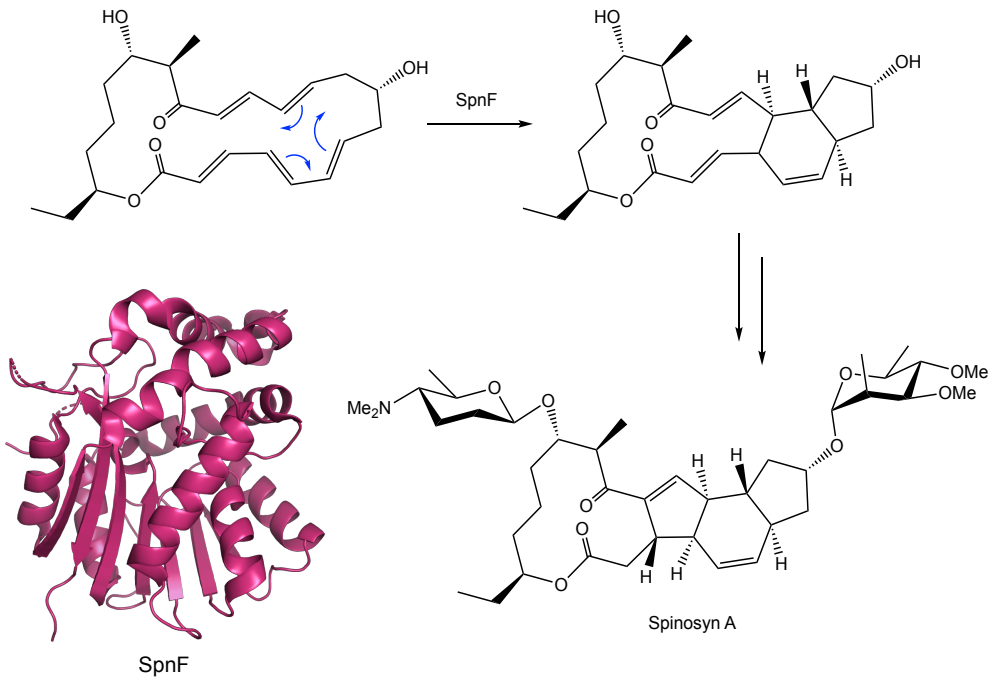


Figure 1.29 Apparent [4+2] cycloaddition catalysed by SpnF in spinosyn biosynthesis. The standalone protein SpnF was shown to catalyse the apparent [4+2] cycloaddition to form the cyclohexene ring of spinosyn[149]. SpnF PBD: 4PNE[155].

1.7.2 VstJ

As alluded to earlier, the spirotetronate moiety of spirotetronate natural products are predicted to form via an intramolecular [4+2] cycloaddition between the exocyclic double bond of the tetronate and 1,3 diene on the terminal end of the polyketide[109]. Analysis of the PKS genes from numerous spirotetronate pathways indicated that all form a 1,3 diene moiety at the appropriate position on their terminal end to react via a [4+2] cycloaddition with the exocyclic double bond on the tetronate ring[110,113,116]. The mechanism of spirotetronate formation was finally solved by studying versipelostatin biosynthesis. Versipelostatin is a spirotetronate produced by a type I modular PKS in *Streptomyces versipellis* 4083-SVS6. This compound downregulates the transcription of *grp78*, an endoplasmic reticulum chaperone protein implicated in Alzheimer's and Parkinson's disease[156]. Within the 110 kbp BGC of versipelostatin a gene encoding a small (142 amino

acids) protein, VstJ, of unknown function was identified[157]. An *S. versipellis* mutant containing a deletion in *vstJ* no longer produced versipelostatin, demonstrating that *vstJ* is an essential biosynthetic gene[157]. Instead, the *S. versipellis* ΔvstJ mutant produced a versipelostatin analogue that was characterised as an apparent pre-[4+2] cyclisation intermediate (**Figure 1.30**). When this precyclisation intermediate was incubated with purified VstJ the spirotetronate moiety formed, demonstrating that VstJ was indeed an apparent [4+2] cyclase responsible for catalysing spirotetronate formation.

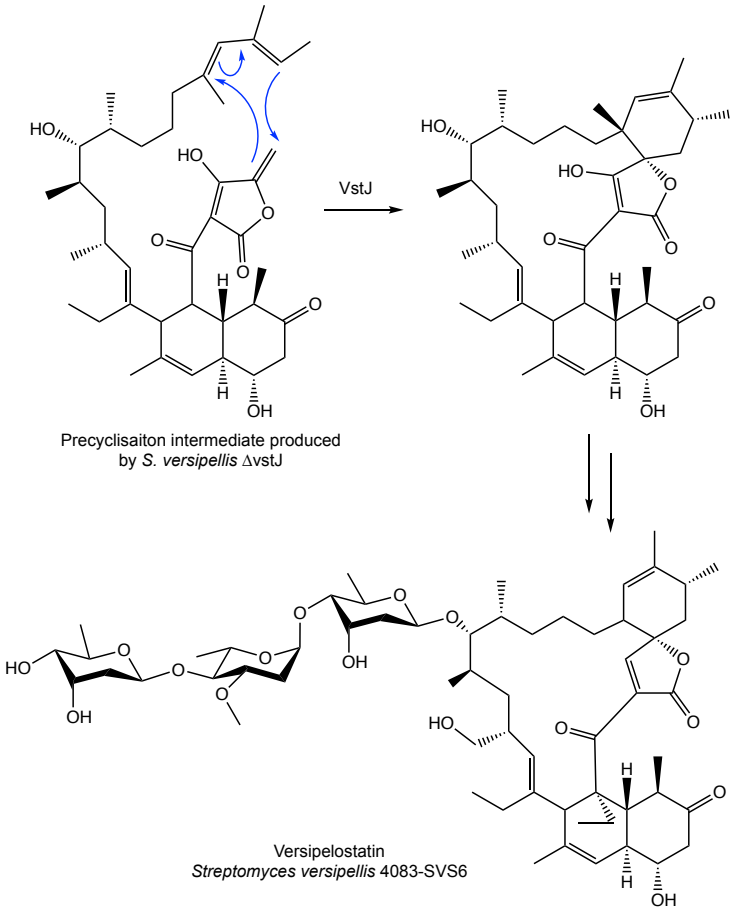


Figure 1.30 Spirotetronate formation in versipelostatin biosynthesis. A single enzyme, VstJ, was shown to catalyse formation of the spirotetronate moiety of versipelostatin[157].

Unlike the previous [4+2] cyclases described, spirotetronate formation did not occur in the absence of VstJ[157]. Consistently, homologues of *vstJ* were identified in the BGCs of other spirotetronates including abyssomicin, chlorothricin, and kijanimicin[157], indicating a conserved role of these enzymes in spirotetronate formation. The structure of the VstJ homologue from the abyssomicin biosynthesis pathway, AbyU, has since been solved and

its activity verified *in vitro*[158] (**Figure 1.31**). AbyU is a homodimer, each monomer consisting of an eight-strand β-barrel structure with an internal active-site cavity. Molecular dynamics simulations were used to dock the linear pre-[4+2] cycloaddition abyssomicin substrate into AbyU. The results from these simulations suggest that the primary role of AbyU is to promote the substrate forming a reactive conformation, rather than provide key catalytic amino acids. The molecular dynamic simulations also indicated that the reaction is indeed concerted, as required for a true pericyclic reaction[158].

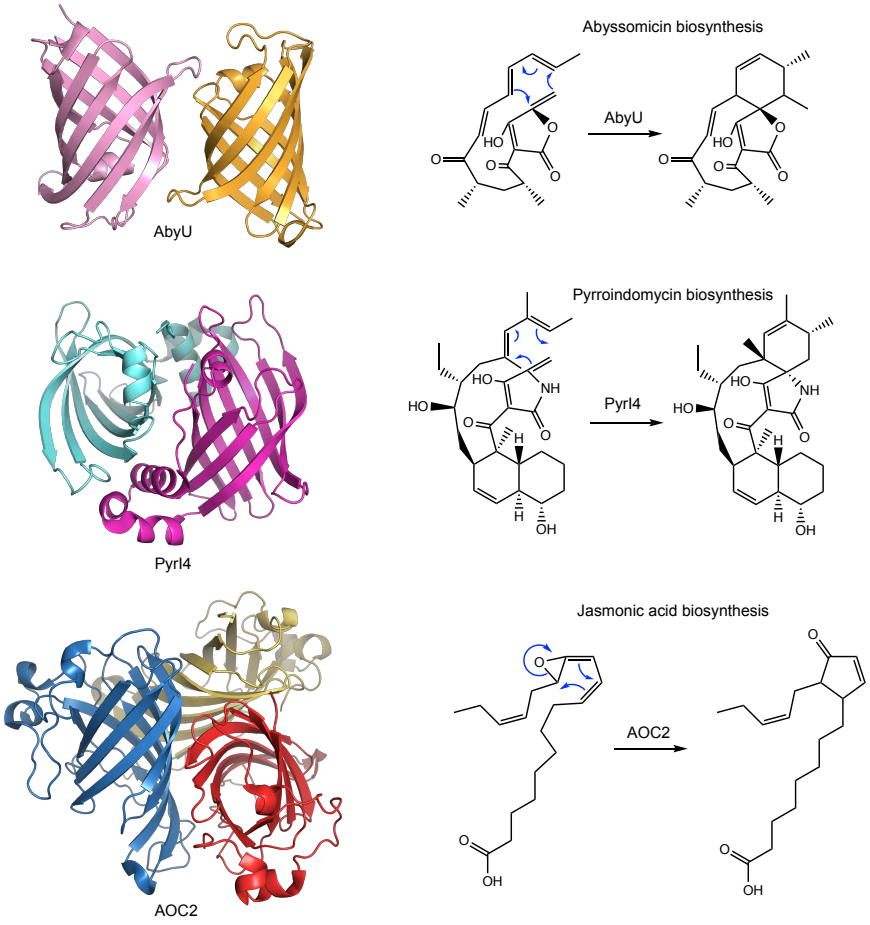


Figure 1.31 The crystal structure of AbyU, Pyrl4, and AOC2. The [4+2] cyclases AbyU[158] (PBD: 5DYQ) and Pyrl4[159] (PDB: 5BTU) share the same fold as AOC2[160] (2DIO). AOC2 catalyses a 4π pericyclic reaction during jasmonic acid biosynthesis.

Another VstJ homologue, Pyrl4, was found to catalyse an equivalent reaction in pyrroindomycin biosynthesis[161]. Pyrroindomycin is a spirotetramate (nitrogen-containing five membered ring) antibiotic produced by *Streptomyces rugosporus* (**Figure 1.30**)[162]. Although Pyrl4 and AbyU only have 24% sequence identity, a crystal structure of Pyrl4

revealed it shares the same eight-strand β-barrel fold[159]. Intriguingly, the same fold is also present in the plant allene oxide cyclase enzymes that catalyse a 4π pericyclic reaction in jasmonic acid biosynthesis (**Figure 1.31**)[158–160]. Like AbyU, the active site of PyrI4 was proposed to function mainly as a hydrophobic binding pocket that imposes conformational constraints of the substrate, making it adopt a reactive conformation[159]. PyrI4 also contains an *N*-terminal domain that, upon substrate binding, adopts an α-helical conformation to act as an essential "lid" domain, holding the substrate in the β-barrel cavity[159] (**Figure 1.32**).

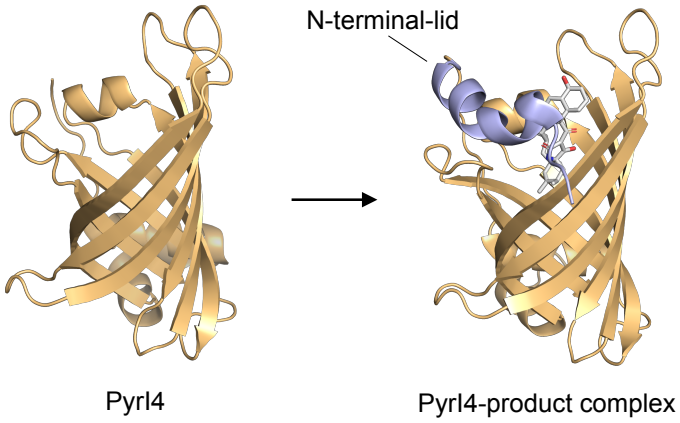


Figure 1.32 *N*-terminal lid domain of the [4+2] cyclase PyrI4. PyrI4 was cocrystallised with its product in its active site, revealing an *N*-terminal lid domain essential for enzyme active[159]. (PDB: 5BTU, 5BU3).

1.7.3 PyrE3

In addition to containing a spirotetronate or spirotetramate moiety, polyketides such as pyrroindomycin, versipelostatin, and chlorothricin also contain a dialkyldecalin ring[109]. The dialkyldecalin was also predicted to form by an intramolecular [4+2] cycloaddition[113,116,161]. Within the 100 kbp BGC of pyrroindomycin a gene, *pyrE3,* was identified that resembles an FAD-dependent monooxygenase. The role of PyrE3 was a mystery, as no monooxygenase was predicted to be required for pyrroindomycin biosynthesis[161]. However, deletion of *pyrE3* in *S. rugosporus* abolished pyrroindomycin production, demonstrating that it is an essential biosynthetic gene. Furthermore, the *S. rugosporus* ΔpyrE3 mutant produced an apparent pyrroindomycin precyclisation substrate lacking both the dialkyldecalin and

spirotetramate moieties. Incubation of this substrate with Tsn11 resulted in formation of the dialkyldecalin, indicating that PyrE3 is really a [4+2] cyclase, not a monooxygenase. The reaction did not proceed in the absence of PyrE3, and the non-covalently bound FAD cofactor was found to play a structural, rather than redox, role in catalysis[161,163]. Following formation of the dialkyldecalin moiety by PyrE3, PyrI4 could then form the spirotetramate ring, as discussed above[161] (**Figure 1.33**). Homologues of PyrE3 were found in the BGCs of all spirotetronates that contain a dialkyldecalin ring, indicating a conserved mechanism of dialkyldecalin formation[157,161,163]. No essential catalytic amino acids were identified in the putative active site of Tsn11, suggesting that, like PyrI4, the role of the active site is to facilitate a reactive substrate conformation[163].

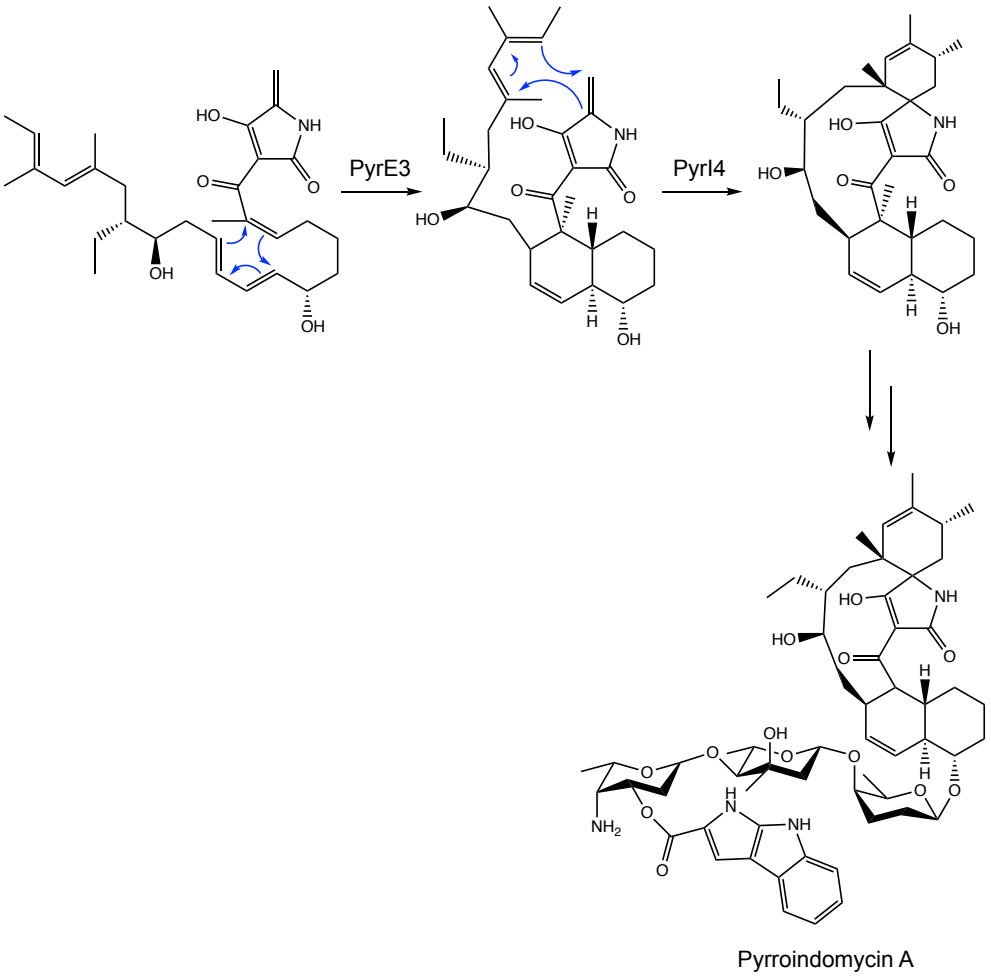


Figure 1.33 Enzymatic [4+2] cyclase cascade in pyrroindomycin biosynthesis. Formation of the dialkyldecalin moiety of pyrroindomycin is catalysed by PyrE3, a [4+2] cyclase that resembles an FAD-dependent monooxygenase. Once the dialkyldecalin moiety has formed PyrI4, a VstJ homologue, catalyses another [4+2] cycloaddition to form the spirotetramate moiety[161]. The same enzymatic cascade is also conserved in spirotetronate biosynthesis.

An interesting recent example of a PyrE3 homologue acting in a non-spirotetronate pathway is found in the biosynthesis of the unusual tetronate tetrodecamycin. Tetrodecamycin contains a dialkyldecalin moiety like pyrroindomycin but is not a spirotetronate[164]. Accordingly, the BGC of tetrodecamycin only encodes a PyrE3 homologue, TedJ, and no VstJ homologue (**Figure 1.34**)[118,165].

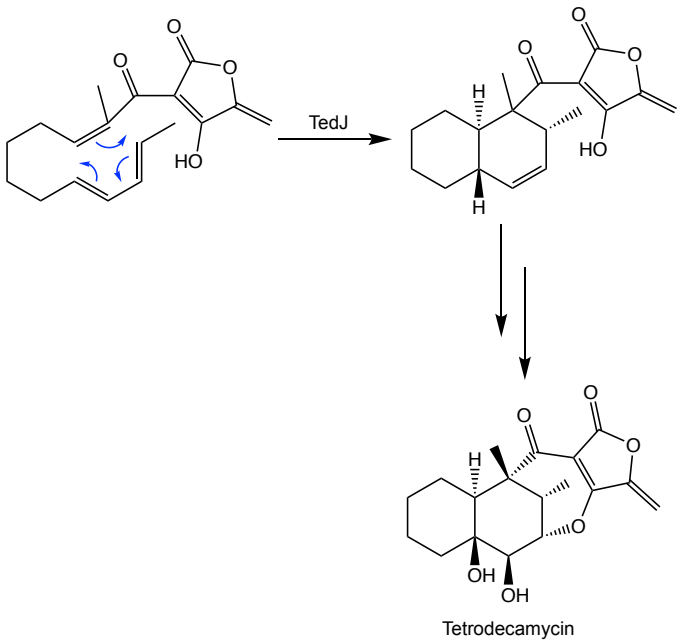


Figure 1.34 The biosynthesis of the dialkyldecalin ring of tetrodecamycin. A PyrE3 homologue TedJ is responsible for catalysing formation of the dialkyldecalin core of tetrodecamycin[165].

1.7.4 The mystery of Tmn8

Unexpectedly, a *vstJ* homologue, *tmn8,* was also identified in the BGC of tetronomycin[157]. Tetronomycin is a polyether tetronate and does not contain a spirotetronate moiety, so the role of Tmn8 is a mystery[157]. Until recently, the polyether tetronate family consisted of three closely related members: tetronomycin[166], tetronasin[167], and SF2487[168]/A80577[169] (referred to in this thesis as tetromadurin). All three are characterised by having four different ring types: a tetronate, a cyclohexane, a tetrahydropyran, and at least one tetrahydrofuran ring (**Figure 1.33**). Tetronomycin and tetronasin are an interesting pair, as they share near-identical carbon skeletons but they have opposite configuration at all ten of their equivalent chiral centers. Recently two new, nearly identical, polyether tetronates were discovered:

nonthmicin[170] and ecteinamycin[171] (**Figure 1.35**), both of which have a tetronate, a tetrahydrofuran and two tetrahydropyran rings. All polyether tetronates are ionophore antibiotics that bind to metal ions to disrupt ion gradients across cell membranes[170–172]. Polyether ionophores are themselves a diverse class of natural product, some of which with notable anticancer activities[173]. For the purposes of this thesis, the term polyether tetronate will only be used to refer to tetronomycin, tetronasin, and tetromadurin (SF2487). Due to their impressive array of different rings, there has been much research on their biosynthesis. Despite this, the mechanism of cyclohexane and tetrahydropyran formation has remained elusive[111,174–177]. The discovery of *tmn8,* a previously unannotated cyclase gene in the tetronomycin BGC offered the possibility that, despite its resemblance to authentic [4+2] cyclases, Tmn8 is actually responsible for creating the cyclohexane and/or tetrahydropyran ring of tetronomycin.

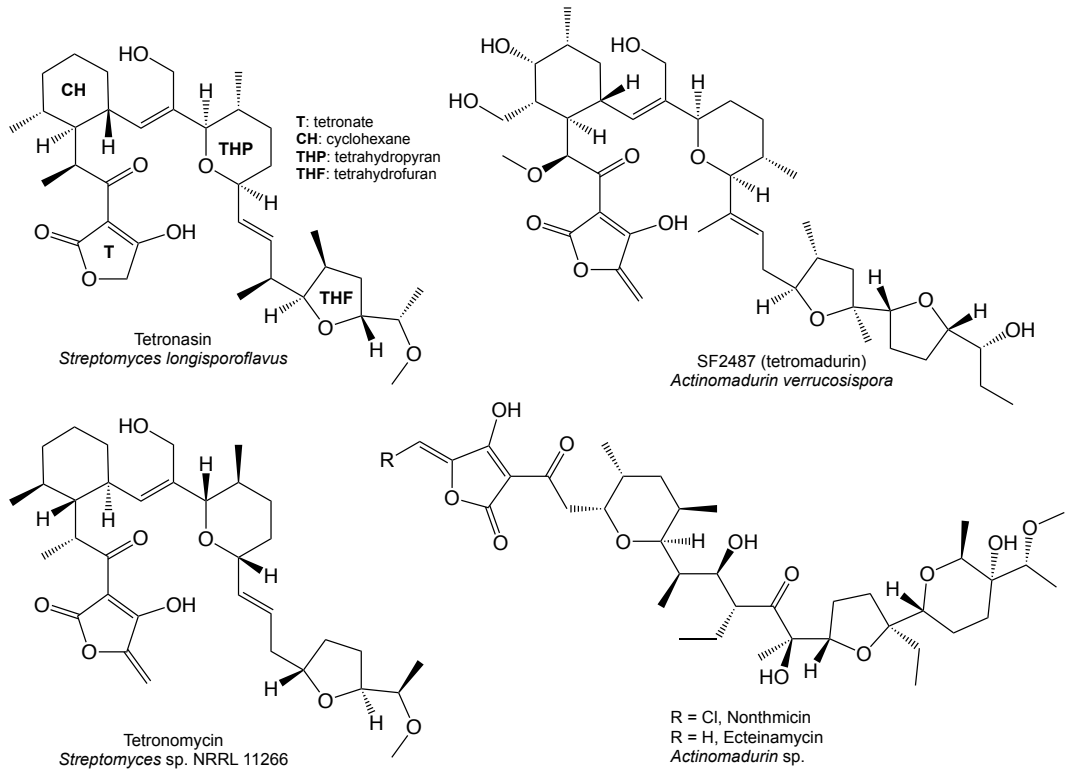


Figure 1.35 Structures of the polyether tetronate ionophore antibiotics. There are five known polyether tetronates, tetronasin, tetronomycin, SF2487/A80577 (tetromadurin), nonthmicin, and ecteinamycin. The different types of ring present in polyether tetronate structures are highlighted on the tetronasin structure. Throughout this thesis the moniker polyether tetronate is only used to designate tetronasin, tetronomycin, and tetromadurin.

1.8 Polyether tetronate biosynthesis

1.8.1 Tetronasin

The biosynthesis of tetronasin (M139603) has been studied more extensively than the other polyether tetronates. Tetronasin is a type I polyketide ionophore antibiotic produced by *Streptomyces longisporoflavus* and was the first polyether tetronate to be discovered; first described in 1981 by ICI Pharmaceuticals (a precursor company of AstraZeneca)[167]. While tetrahydrofuran and tetrahydropyran rings are common in other ionophores, such as monensin[178] and salinomycin[179], owing to their cation binding ability, tetronasin was the first example of a polyether ionophore to contain a cyclohexane ring. Its unusual structure and antibiotic activity led to numerous partial and chemical synthesis studies[180–184]. Tetronasin can bind to numerous metal ions but forms the most stable complexes with sodium ions (Na^+), which it binds at a 1:1 ratio[185,186]. Bioactivity assays revealed antibiotic activity against numerous ruminal microorganisms, particularly Gram-positive bacteria[187]. When added to feedstock, tetronasin has been shown to enhance weight gain in cows[188] and sheep, also leading to increased wool production in the latter case[189]. Tetronasin has been banned in Europe since 2006, but is still used in animal husbandry in other parts of the world[109].

The first biosynthetic studies on tetronasin began in the 1980s, before the modular nature of type I PKS enzymes was known. The primary areas of interest at the time were in identifying its biosynthetic precursors and probing the chemistry of formation of the four different rings from a hypothetical linear tetronasin intermediate (**Figure 1.36**).

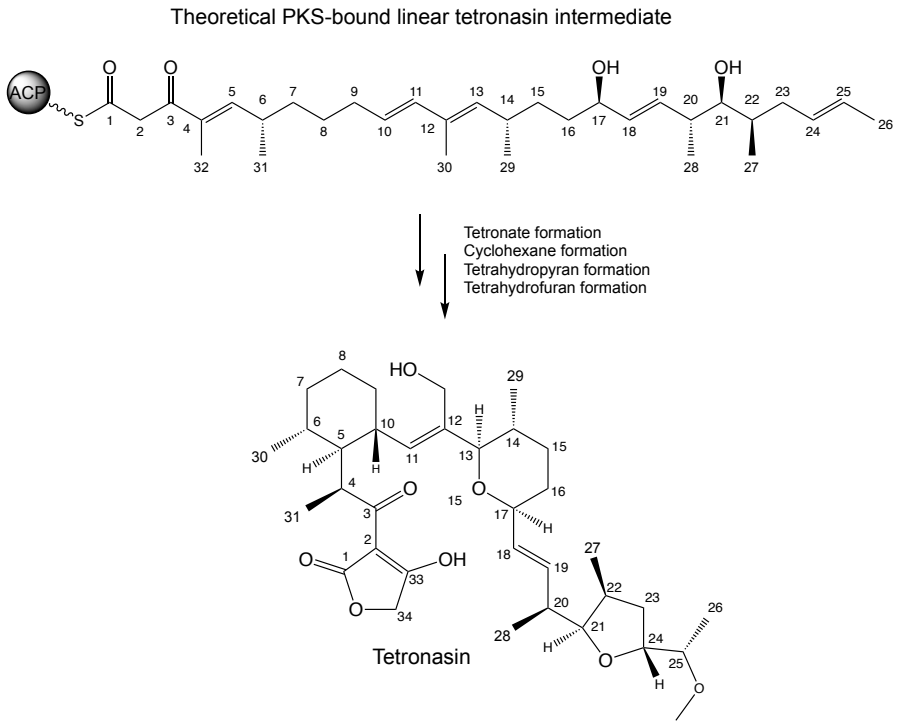


Figure 1.36 Formation of tetronasin from a linear intermediate. Polyketide synthase enzymes (PKS) were predicted to make the displayed linear tetronasin intermediate. From this intermediate four different rings are formed: a tetronate, cyclohexane, tetrahydropyran, and tetrahydrofuran[177].

The biosynthetic precursors of tetronasin were identified by isotope feeding experiments. [1-^{13}C] acetate, [1,2-^{13}C] acetate, and [1-^{13}C] propionate were fed to a *S. longisporoflavus* culture and the site of isotope incorporation was determined using ^{13}C-^{13}C NMR. Once taken into the cell the isotopically labelled acetate and propionate precursors were converted into malonyl-CoA and (*2S*)-methylmalonyl-CoA forms, respectively, enabling them to be used as PKS extension units[61]. The feeding studies demonstrated that the carbon backbone of tetronasin is comprised of six propionate units and seven acetate units[174] (**Figure 1.37**). Feeding [Me-C^{13}] methionine additionally demonstrated that the methyl of the C25 methoxy group is derived from methionine (in the form of S-adenosyl methionine)[174]. Notably, two of the carbon atoms of the tetronate ring (C33 and C34) were not labelled. It was established several decades later that these carbons are derived from hydroxypyruvate, converted into a glycolyl-ACP by a dedicated multienzyme complex patterned on pyruvate dehydrogenase enzymes (*Sun et al, unpublished*) (**Figure 1.37**).

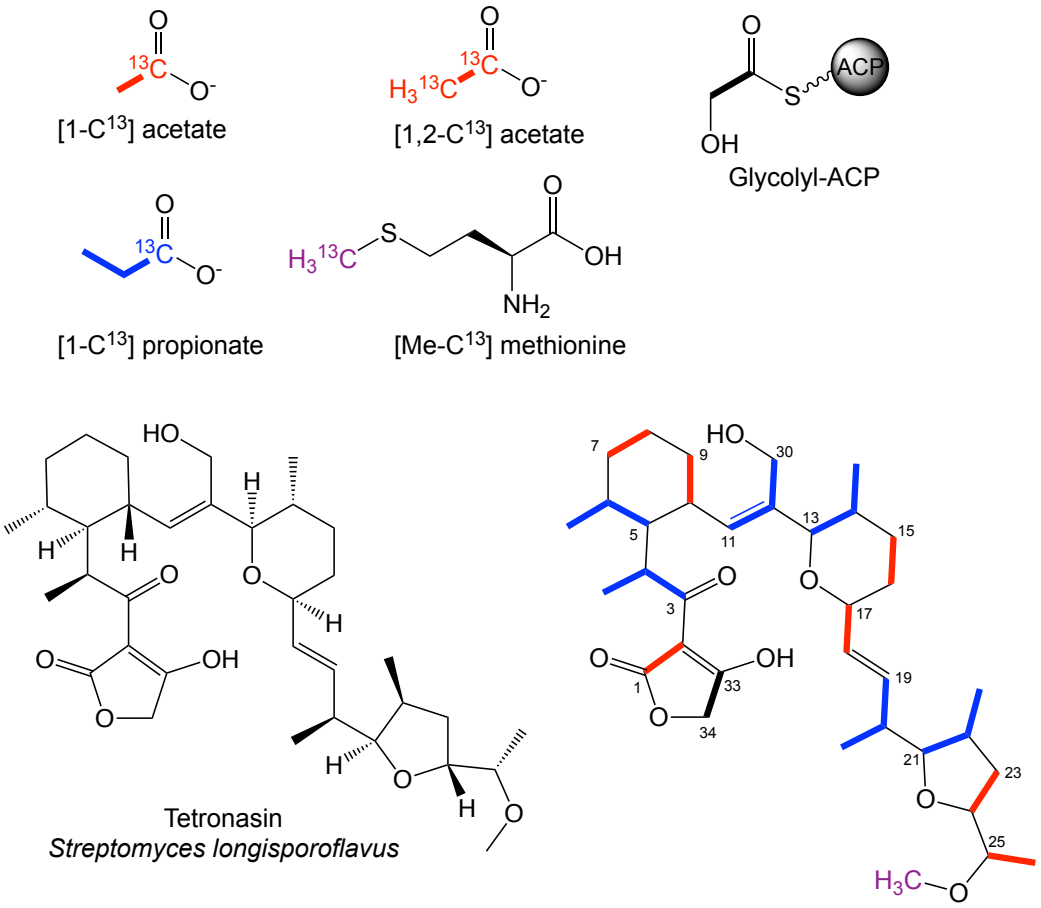


Figure 1.37 Summary of the ^{13}C feeding experiments in *S. longisporoflavus*. [1-C^{13}] acetate, [1,2-C^{13}] acetate, [1-C^{13}] propionate, and [Me-C^{13}] methionine were fed to a growing culture of *S. longisporoflavus* and their incorporation detected using NMR[174]. Several decades later the C33 and C34 carbons of the tetronate ring were discovered to be derived from glycolyl-ACP.

The early isotopic precursor experiments also provided insight into the mechanism of tetrahydrofuran formation. Feeding *S. longisporoflavus* with [1-C^{13}, 1-$^{18}O_2$] acetate and [1-C^{13}, 1-$^{18}O_2$] propionate revealed no incorporation of ^{18}O in the oxygen atoms at C25 or C30, indicating that they are not derived from the fed acetate or propionate units[191] (**Figure 1.38**).

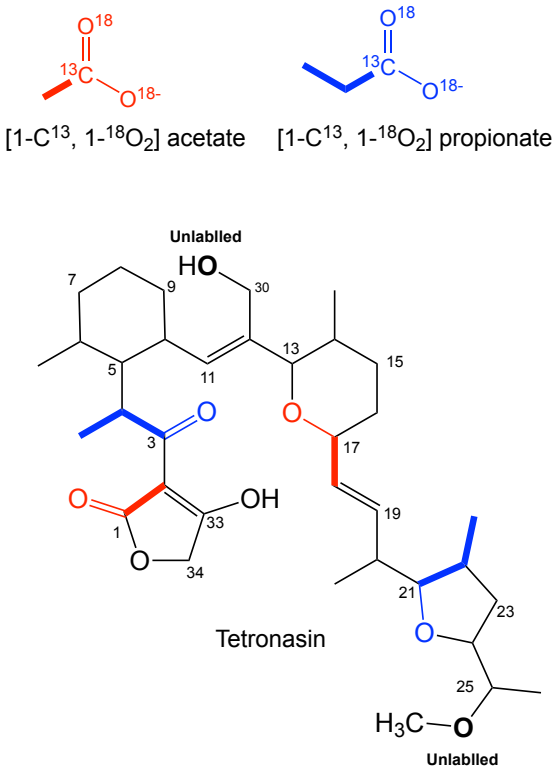


Figure 1.38 Summary of ^{18}O feeding experiments in *S. longisporoflavus*. [1-C^{13}, 1-$^{18}O_2$] acetate and [1-C^{13}, 1-$^{18}O_2$] were fed to *S. longisporoflavus* and the incorporation of ^{18}O was detected using NMR[175]. All of the oxygen atoms in tetronasin except that of the C25 methoxy group and the C30 acetate group are derived from an acetate or propionate extender unit.

The lack of ^{18}O incorporation in C25 suggested that the tetrahydrofuran formed by oxidative cyclisation, a common mechanism of tetrahydrofuran formation in natural products. Oxidative cyclisation involves the oxidation of an *E*-alkene to a reactive epoxide that is attacked by an adjacent hydroxyl to form a tetrahydrofuran ring[190,191] (**Figure 1.39a**). In such cases, epoxide formation is regio- and stereospecifically catalysed by an epoxidase enzyme, followed by epoxide hydrolysis by an epoxide hydrolase to form the tetrahydrofuran ring[191–194]. In the case of tetronasin, an *E*-double bond would have to initially be present between C24-C25 before being stereospecifically oxidised to an epoxide. A hydroxyl group present at C21 would then need to attack C24 to form the tetrahydrofuran ring and the unlabelled C25 hydroxyl (derived from molecular oxygen). Feeding *S. longisporoflavus* isotopically labeled tri- and tetraketide tetronasin intermediates containing the C21 hydroxyl and the C24-C25 *E*-alkene showed that they were incorporated intact into tetronasin, providing strong support for this mechanism[177,195] (**Figure 1.39b**).

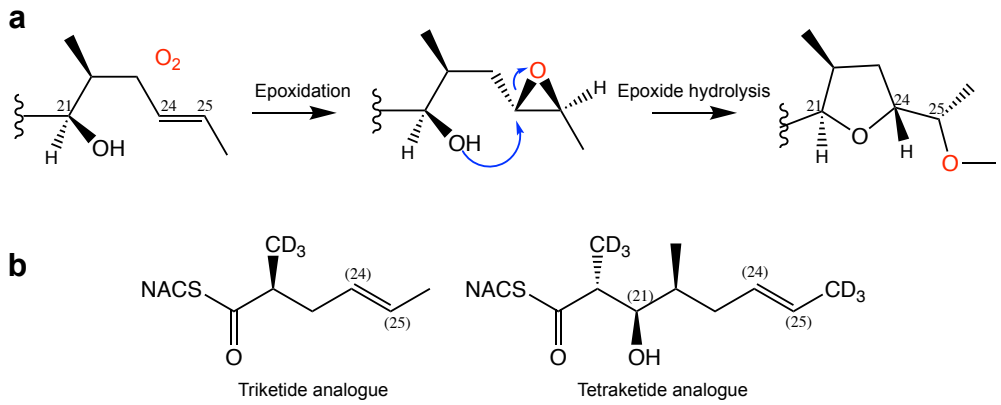


Figure 1.39 Proposed mechanism of tetrahydrofuran formation in tetronasin biosynthesis. a, The tetrahydrofuran ring of tetronasin forms via an oxidative cyclisation mechanism of epoxidation and epoxide hydrolysis[175]. **b,** The isotopically labelled tetronasin di- and triketide analogues that were incorporated into tetronasin when fed to *S. longisporoflavus*[177,195]. The numbers is brackets correspond to the atom number in the final tetronasin structure.

The feeding studies with [1-^{13}C, 1-$^{18}O_2$] acetate/propionate precursors also provided some of the first clues as to how the tetrahydropyran and cyclohexane rings might form. The oxygen at C17 was labelled by feeding [1-^{13}C, 1-$^{18}O_2$] acetate, indicating that the tetrahydropyran ring likely forms by the C17 hydroxyl attacking C13[175]. Numerous mechanisms were proposed by which the electrophilicity of C13 might be enhanced to promote nucleophilic attack by the C17 hydroxyl[175,176]. One of the mechanisms resonated more than the others: a scheme in which both the tetrahydropyran and cyclohexane rings are closed in a single step, after the tetronate and tetrahydrofuran rings have already formed (in a compound labelled pretetronasin) (**Figure 1.40**). Stereochemical and electronic analysis of a pretetronasin model indicated that it could theoretically fold into a low-energy conformation where groups that needed to form bonds to make tetronasin were within bonding distance[175,177].

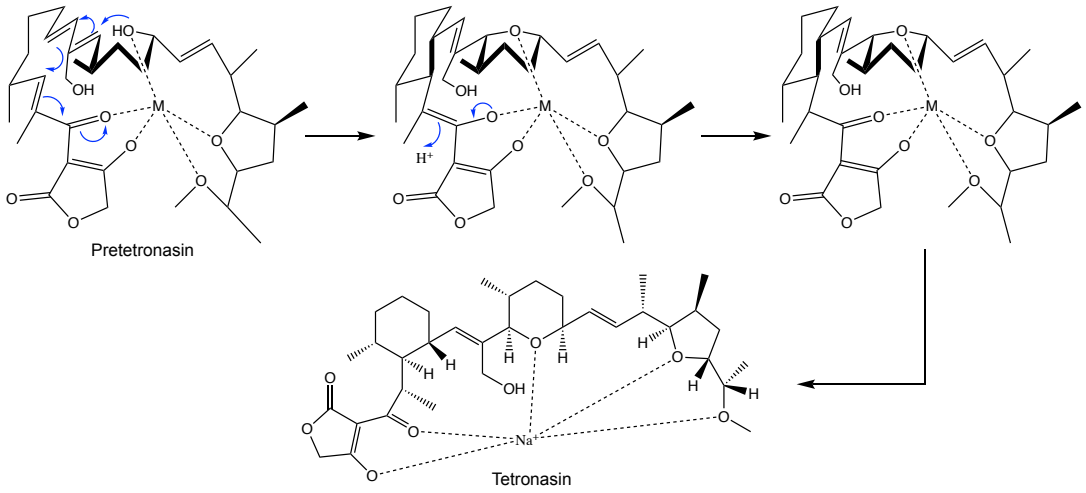


Figure 1.40 Metal-catalysed mechanism of cyclohexane and tetrahydropyran formation in tetronasin biosynthesis. One of the proposed mechanisms for formation of the cyclohexane and tetrahydropyran rings in tetronasin is a metal-catalysed ring-closing cascade[175,177].

It was further speculated that such a reactive conformation could be promoted by pretetronasin complexing to a metal ion such as Na^+, Mg^{2+}, or Zn^{2+} [175,177]. The proposed active metal-bound conformation of pretetronasin even resembles the final conformation of tetronasin in solution, providing additional support for this mechanism[167,177]. While this mechanism was never demonstrated *in vitro* or *in vivo*, an analogous reaction was performed during the second published total synthesis of tetronasin (**Figure 1.41**)[180]. Using a substrate analogous to pretetronasin, the cyclohexane and tetrahydropyran rings could be formed in a single step, with the correct stereochemistry, using a basic catalyst[180]. It is therefore possible that tetrahydropyran and cyclohexane formation in polyether tetronates does not require an enzyme[175].

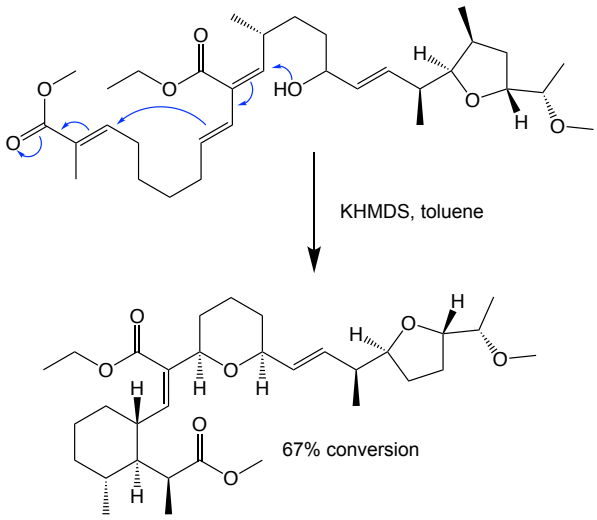


Figure 1.41 Method used for cyclohexane and tetrahydropyran in tetronasin total synthesis. The reaction used potassium bis(trimethylsilyl)amide (KHMDS) in toluene to produce a single product at 67% yield[180].

In 2007 the BGC of tetronasin (the *tsn* cluster) was cloned and sequenced by the Leadlay lab (GenBank: FJ462704). The *tsn* gene cluster is roughly 125 kbp in size and contains ca. 38 genes (**Figure 1.42** and **Table 1.2**).

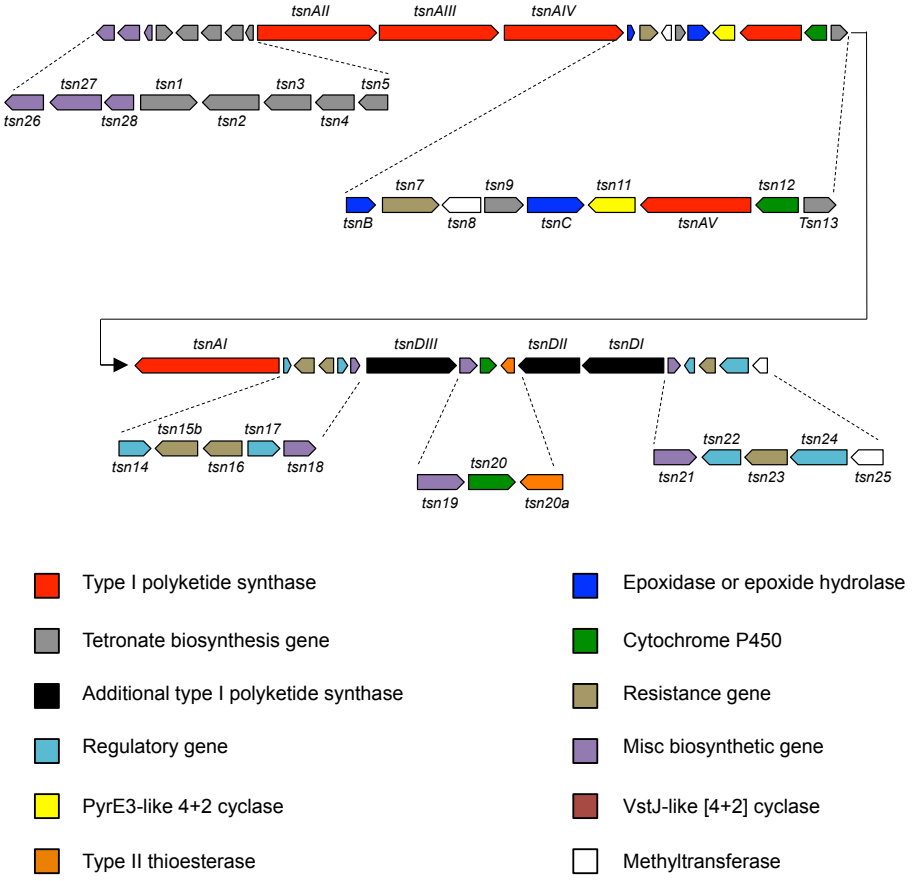


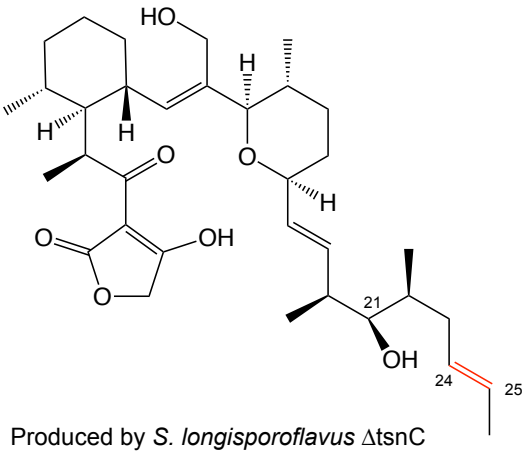
Figure 1.42 The biosynthetic gene cluster of tetronasin. The tetronasin gene cluster is 125 kbp in size. Genes are not drawn to scale. Genes are colour coded according to their predicted functions.

Table 1.2 Table 1.2 Tetronasin gene cluster annotation table.

orf	Size	Protein homologue and origin	Identity/ similarity (%)	Proposed function
tsn26	218	WP_078074012, *Streptomyces niveus*	96/97	4'-phosphopantetheinyl transferase
tsn27	293	WP_078074014, *Streptomyces niveus*	97/97	Metallophosphoesterase
tsn28	145	WP_078079323, *Streptomyces niveus*	92/97	Glyoxalase
tsn1	272	WP_100303314, *Streptomyces* sp. CNZ306	65/78	Acyltransferase
tsn2	431	WP_117398905, *Actinomadura* sp. LHW52907	48/59	E2 dihydrolipoyl acyltransferase
tsn3	334	WP_117398904, *Actinomadura* sp. LHW52907	70/80	Pyruvate dehydrogenase E1 beta subunit
tsn4	345	WP_117398903, *Actinomadura* sp. LHW52907	73/79	Pyruvate dehydrogenase E1 alpha subunit
tsn5	108	WP_117398902, *Actinomadura* sp. LHW52907	54/69	Acyl-carrier protein
tsnAII	5136	WP_076682134, *Streptomyces alfalfae*	50/61	Type I polyketide synthase: (KS, AT, DH, KR, ACP, KS, AT, DH, KR, ACP, KS, AT, KR, ACP)
tsnAIII	5630	BAE93730 (TmnAIV), *Streptomyces* sp. NRRL 11266	53/64	Type I polyketide synthase: (KR, AT, DH, ER, KR, ACP, KS, AT, DH, KR, ACP, KS, AT, DH, KR, ACP)
tsnAIV	6036	WP_047014291, *Streptomyces* sp. CNQ-509	52/63	Type 1 polyketide synthase: (KS, AT, DH, ER, KR, ACP, KS, AT, DH, ER, KR, ACP, KS, AT, DH, KR, ACP)
tsnB	139	CAQ64695, *Streptomyces lasaliensis*	46/57	Epoxide hydrolase
tsn7	422	WP_076474008, *Micromonospora avicenniae*	58/73	MFS transporter
tsn8	290	WP_027756739, *Streptomyces* sp. CNH099	49/68	Methyltransferase
tsn9	350	WP_030750290, *Streptomyces* sp. NRRL S-31	56/70	FabH 3-oxoacyl-ACP synthase III
tsnC	477	B5M9L6 (LasC), *Streptomyces lasaliensis*	53/67	Epoxidase
tsn11	495	WP_100249299, *Streptomyces tsukubensis*	46/60	[4+2] cyclase
tsnAV	1553	WP_066029228, *Streptomyces hygroscopicus*	51/63	Type I polyketide synthase (KS, AT, KR, ACP)
tsn12	394	KWX00849, *Streptomyces thermoautotrophicus*	53/68	Cytochrome P450
tsn13	350	WP_099855277, *Micromonospora* sp. CNZ299	56/71	FabH 3-oxoacyl-ACP synthase
tsn15	206	BAF73716 (Tmn8), *Streptomyces* sp. NRRL 11266	29/44	[4+2] cyclase
tsnAI	5029	CAQ64686, *Streptomyces lasaliensis*	52/63	Type I polyketide synthase (KS^Q, AT, ACP, KS, AT, DH, KR, ACP, KS, AT, DH, ER, KR, ACP)
tsn14	178	KDN76189, *Streptomyces olindensis*	59/76	Transcriptional regulator
tsn15b	540	WP_050066454, *Rhodococcus* sp. RD6.2	51/69	Tetronasin ABC transporter
tsn16	301	KDN76193, *Streptomyces olindensis*	78/87	ABC transporter, ATP binding component
tsn17	256	KDN76194, *Streptomyces olindensis*	71/82	Pathway specific activator
tsn18	294	KUJ69111, *Streptomyces albus subsp. albus*	83/89	Dioxygenase
tsnDIII	2097	WP_018851410, *Streptomyces* sp. CNY243	66/75	Type I polyketide synthase
tsn19	520	KUJ69086, *Streptomyces albus subsp. albus*	77/85	Acyl-CoA synthase
tsn20	423	WP_113685098, *Streptomyces* sp. PT12	61/76	Cytochrome P450
tsn20a	255	WP_103534038, *Streptomyces* sp. SM11	67/75	Thioesterase
tsnDII	1502	BAE93739 (TmnDII), *Streptomyces* sp. NRRL 11266	58/67	Type I polyketide synthase (KS, AT, DH, ACP, TE)
tsnDI	2476	WP_100303217, *Streptomyces* sp. CNZ306	62/71	Type I polyketide synthase (AT, ACP, KS, ACP, KS, ACP, KR)
tsn21	585	PJJ38674, *Streptomyces* sp. CNZ306	69/81	Acyltransferase
tsn22	249	WP_037741229, *Streptomyces* sp. CNQ-525	62/71	Transcriptional regulator
tsn23	379	BAE93743 (Tmn22), *Streptomyces* sp. NRRL 11266	56/74	Putative membrane protein
tsn24	918	WP_079053178, *Streptomyces phaeochromogenes*	55/68	LuxR transcriptional regulator
tsn25	249	WP_078074018, *Streptomyces niveus*	92/95	Methyltransferase

The *tsn* cluster contains eight type I PKS enzymes, five of which (*tsnAI-tsnAV*) appear to be involved in the biosynthesis of the 26-carbon backbone of tetronasin. Together, these five PKS enzymes contribute a loading module and 12 extension modules. The polyketide synthase enzymes are responsible for establishing the stereochemistry of the C6, C14,

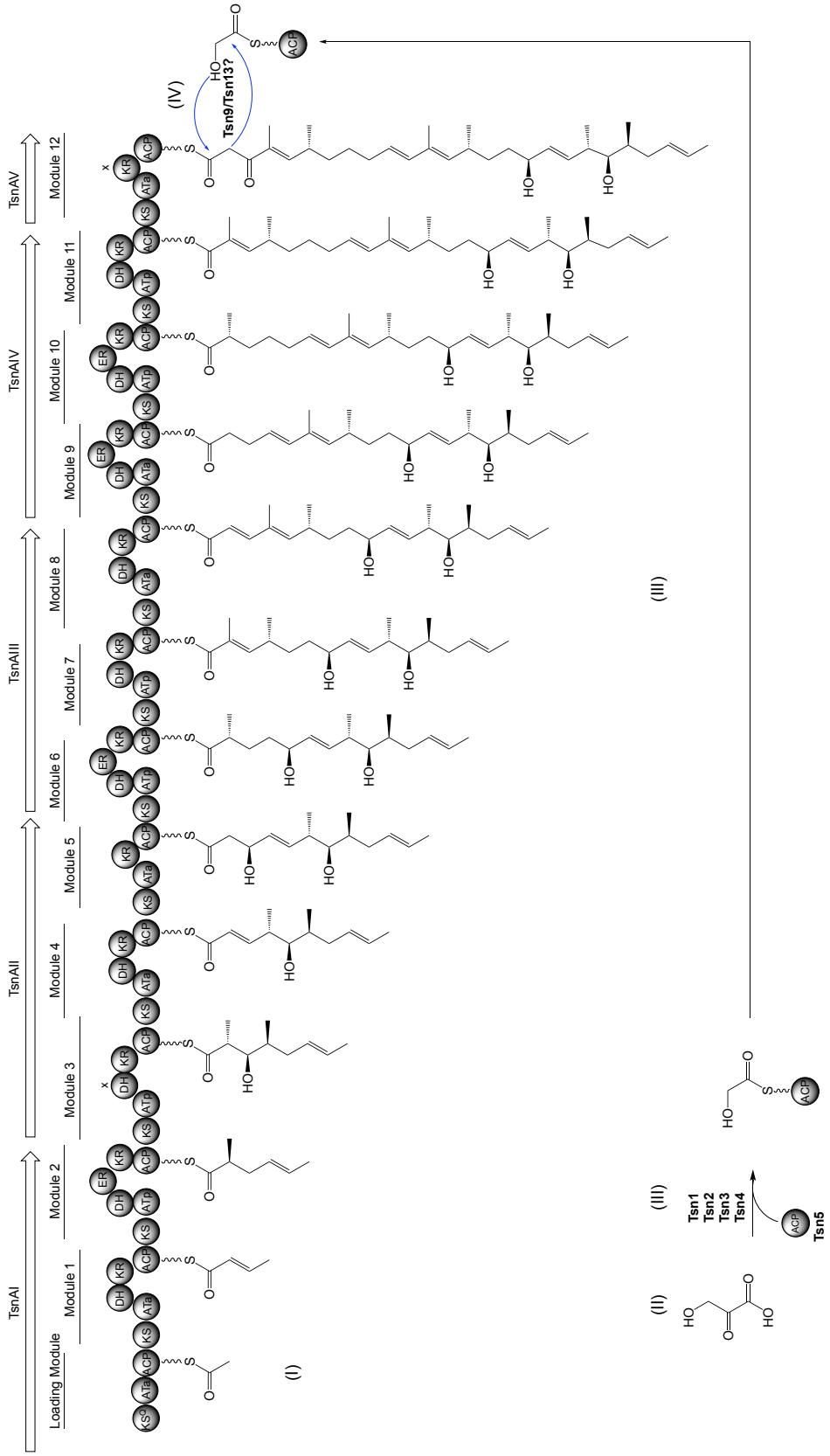
C17, C20, C21, and C22 chiral centers. The configurations of the remaining stereocenters are established during ring formation. At one end of the *tsn* cluster are three additional type I PKS enzymes (*tsnDI-tsnDIII)* that, on the basis of the domains they contain, are very unlikely to be involved in tetronasin biosynthesis. In fact, deletion of *tsnDIII* did not affect tetronasin biosynthesis (Demydchuk *et al*, *unpublished*). The nine genes adjacent to *tsnDI-tsnDIII* (*tsn18-tsn25)* also have no obvious role in tetronasin biosynthesis, and deletion of *tsn18* had no affect on tetronasin biosynthesis. At the other end of the cluster are genes encoding the E1 (*tsn3* and tsn4) and E2 (*tsn3* and *tsn4*) components of the hydroxypyruvate dehydrogenase complex and an acyl carrier protein (*tsn5*). Together these proteins have been shown to produce glycolyl-ACP, the origin of the two additional carbon atoms (C33 and C34) of the tetronasin tetronate ring (Sun *et al*, *unpublished*). Unlike other tetronate BGCs, which only contain a single gene for a FabH-like protein, the *tsn* cluster contains two: *tsn9* and *tsn13*. Interestingly, both FabH-like proteins were shown to be essential for tetronasin biosynthesis by targeted inactivation (Sun *et al*, *unpublished*). The role each enzyme plays in tetronate ring formation/PKS chain release is therefore unknown. The gene *tsn8* encodes a putative methyltransferase enzyme. Selective inactivation of *tsn8* resulted in production of a new metabolite that corresponded in mass to an analogue of tetronasin lacking a methyl group, suggesting it is likely responsible for catalysing formation of the C25 methoxy group. Similarly, inactivation of the P450 monooxygenase *tsn12* resulted in production of a tetronasin analogue lacking a hydroxyl group, indicating it is likely responsible for catalysing hydroxylation of C30 (Sun *et al*, *unpublished*). Also within the *tsn* cluster are the epoxidase (*tsnC*) and epoxide hydrolase (*tsnB*) genes responsible for catalysing stereoselective tetrahydrofuran formation. Deletion of *tsnC* resulted in production of a novel tetronasin metabolite containing all rings except the tetrahydrofuran. In the place of the tetrahydrofuran ring was the predicted C24-C25 *E* alkene (**Figure 1.43**) (Hong *et al*, *unpublished*).



Produced by *S. longisporoflavus* ΔtsnC

Figure 1.43 Tetronasin analogue produced by a *S. longisporoflavus* ΔtsnC. Deletion of *tsnC* in *S. longisporoflavus* resulted in the production of a tetronasin analogue where the tetrahydrofuran ring had not yet formed (structure solved by NMR). It its place was a C24-C25 *E* alkene (Hui Hong, *unpublished*).

In addition to verifying the role of *tsnC* in tetrahydrofuran formation, the accumulation of this intermediate demonstrated that formation of the other rings is not dependent on the tetrahydrofuran ring having formed. Whether the cyclohexane, tetrahydropyran, and tetrahydrofuran rings form while the polyketide is attached to the PKS or after chain release is unknown, though evidence from other polyether pathways suggests that the tetrahydrofuran and tetrahydropyran rings can form before chain release[196,197]. Finally, the *tsn* gene cluster also contains several transcriptional regulator genes (*tsn14* and *tsn17)*, resistance proteins (*tsn15b, tsn16*, and *tsn7)*[198], and a PPTase (*tsn26).* Unfortunately, analysis of the cluster up until 2015 provided no hints as to how the cyclohexane and tetrahydropyran rings formed (the putative cyclases described in this thesis, Tsn11 and Tsn15, had not been identified). At the time, none of the enzymes in the BGC appeared to be cyclases, nor did the PKS enzymes contain any putative pyran synthase domains[68,69], that could account for tetrahydropyran or cyclohexane biosynthesis. The proposed tetronasin biosynthesis, based on the state of the art in 2015, is presented in **Figure 1.44**.


TsnAI
TsnAII
TsnAIII
TsnAIV
TsnAV
Loading Module
Module 1
Module 2
Module 3
Module 4
Module 5
Module 6
Module 7
Module 8
Module 9
Module 10
Module 11
Module 12
(I)
(II)
(III)
(IV)
Tsn1
Tsn2
Tsn3
Tsn4
Tsn5
Tsn9/Tsn13?

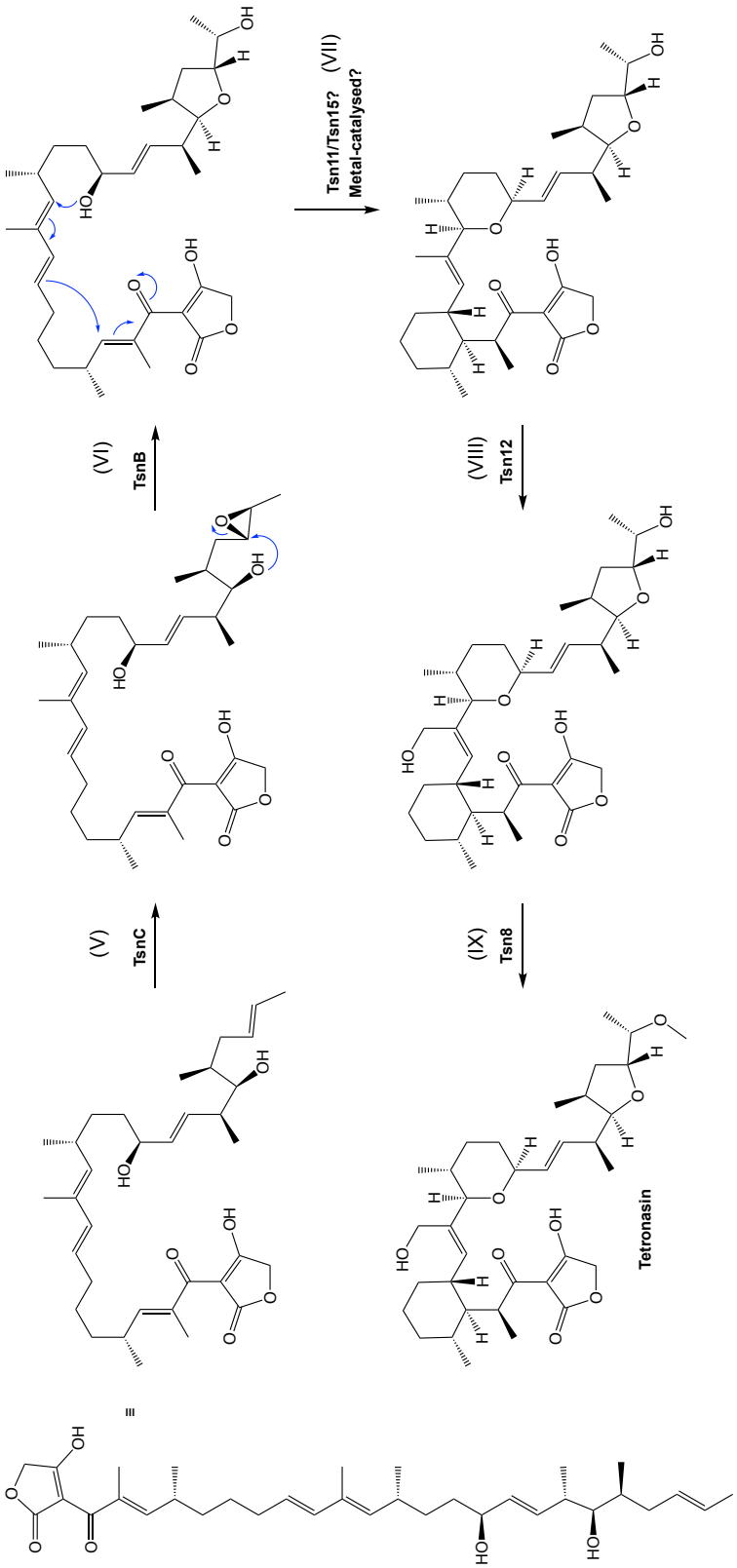

(VII)
Tsn11/Tsn15?
Metal-catalysed?
(VI)
TsnB
(VIII)
Tsn12
(V)
TsnC
(IX)
Tsn8
Tetronasin
III

Figure 1.44 Proposed biosynthetic pathway of tetronasin. The relative timing of the different reactions are unknown and could be different from the order presented. (I) The carbon backbone of tetronasin is synthesised by a TsnAI-A12. (II) Hydroxypyruvate is the origin of the two addition carbons of the tetronate ring. (III) Pyruvate hydroxydehydrogenase enzymes convert hydroxypyruvate into glycolyl-ACP. (VI) The two FabH enzymes in the tetronasin BGC likely both contribute to tetronate ring formation, though the exact role of each is unknown. (V) Epoxidation catalysed by TsnC. (VI) Epoxide hydrolysis catalysed by TsnB. (VII) The timing and mechanism of cyclohexane and tetrahydropyran formation is unknown. (VIII) Tsn12 catalyses hydroxylation of C30. (IX) Tsn8 catalyses methylation of the C25 hydroxyl.

1.8.2 Tetronomycin

The BGC of tetronomycin (the *tmn* cluster) has also been sequenced[111] (**Figure 1.45**). First isolated from the terrestrial *Streptomyces* sp. NRRL 11266, the tetronomycin gene cluster has also been identified in sea sponge-associated actinomycetes[199] and in polar desert soils[200]. Like tetronasin, tetronomycin also readily binds to sodium ions, and has antibiotic activity towards numerous gram-positive bacteria, including *Staphylococcus aureus*[166,201]. Several partial and total chemical syntheses of tetronomycin have been performed[202–205], but only one study has investigated its biosynthesis[111].

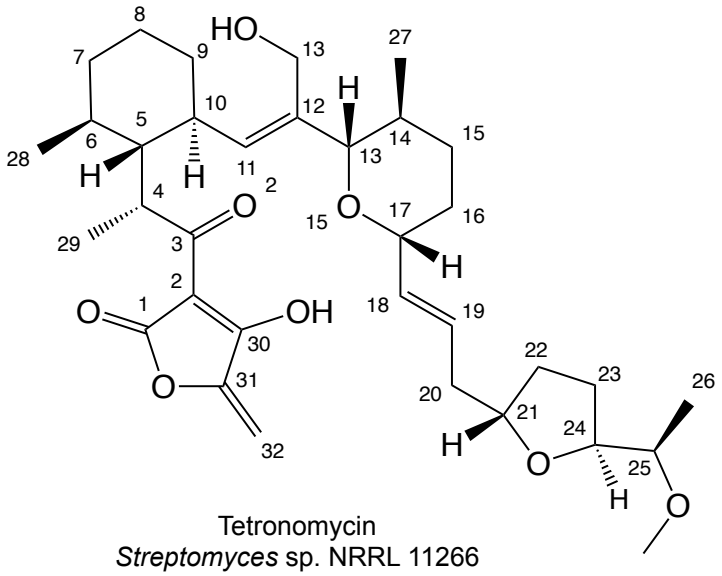


Figure 1.45 Structure and atom numbers of tetronomycin

Given the structural similarities between tetronasin and tetronomycin, it is no surprise that the *tmn* cluster contains similar genes to the *tsn* cluster; however, it also has some notable differences[111]. The 26-carbon backbone of tetronomycin is synthesised by six type I PKS enzymes (TmnAI-TmnAVI), together contributing a loading module and 12 extension modules[111]. The PKS enzymes are responsible establishing the stereochemistry of C6, C14, C17, and C21 in the opposite configuration to the equivalent centres in tetronasin (**Figure 1.46**). The other differences in stereochemistry must therefore be introduced during formation of the tetrahydrofuran, tetrahydropyran and cyclohexane rings.

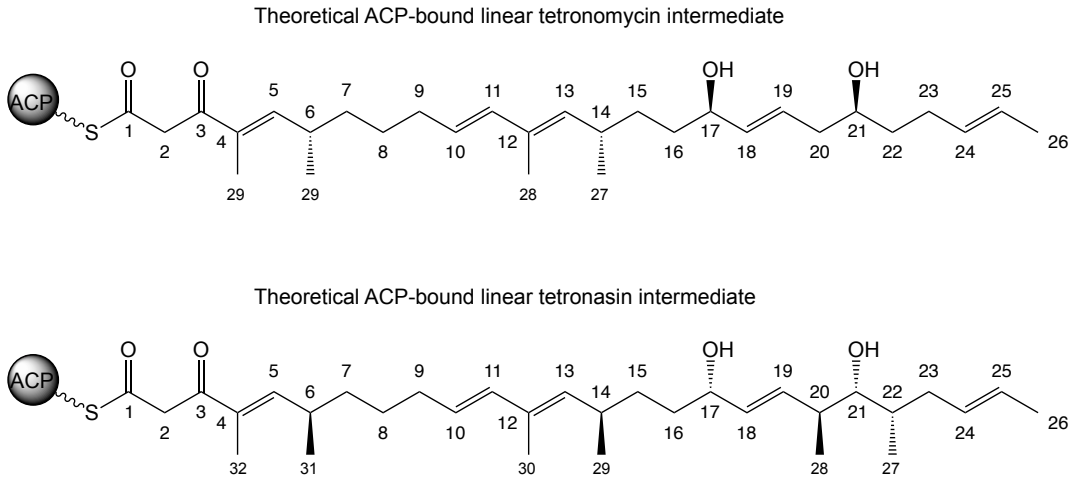


Figure 1.46 Theoretical linear intermediate of tetronomycin. The stereochemistry at C6, C14, C17, C21 is introduced by the *tmn* I PKS enzymes. The theoretical linear tetronasin intermediate is included for the sake of comparison.

Interestingly, like the *tsn* cluster the *tmn* cluster also contains additional type I PKS genes at one end with no obvious function (analogous to the *tsnD* genes identified in the *tsn* cluster). As in the case of the *tsn* cluster, deletion of one of these PKS genes (*tsnDI)* confirmed its lack of involvement in tetronomycin biosynthesis. The tetronate ring of tetronomycin is synthesised in the manner typical for tetronates, using glyceryl-ACP (the origin of C31, C32, and C33)[109,111] (**Figure 1.41**). The substrate 1,3-bisphosphoglycerate is converted into glyceryl-ACP by the FkbH-like protein Tmn16 and the acyl-carrier protein Tmn7a[111]. A single FabH-like protein, Tmn15 then catalyses C-C and C-O bond formation between glyceryl-ACP and the PKS-bound tetronomycin acyl chain, producing the tetronate ring and causing chain release[111,122]. *agg4* (tmn20) and *agg5* (tmn18) homologues are present for the formation of the exocyclic double bond on the tetronate (not present in tetronasin)[111,123]. Like tetronasin, one of the methyl groups of tetronomycin is hydroxylated (C28) and it has a methoxy group at C25. Correspondingly, the *tsn* cluster contains a cytochrome P450 (Tmn14) and a methyltransferase (Tmn12), respectively responsible for the hydroxylation and methylation[111]. Like the *tsn* cluster, the *tmn* cluster also contains the epoxidase (*tmnC*) and epoxide hydrolase (*tmnB*) genes required for stereoselective tetrahydrofuran formation via oxidative cyclisation. In regards to the cyclohexane and tetrahydropyran rings, the authors proposed that they formed simultaneously in a reaction very similar to the metal-catalysed ring closure reaction originally proposed for tetronasin[111]. However it was also hypothesised that an enzyme

could be responsible. The discovery of the [4+2] cyclase homologue, *tmn8,* in the *tmn* cluster adds fresh impetus to this hypothesis.

1.9 Aims of this thesis

The central aim of this thesis was to deconvolute the biosynthetic pathway to polyether tetronates, a particularly intriguing and unsolved problem in understanding how initial PKS products are rearranged into antibiotic products. A particular aim was to examine the role of putative [4+2] cyclase homologues recently discovered in the biosynthetic gene clusters of polyether tetronates, and to determine whether these enzymes played a role in tetrahydropyran and/or cyclohexane biosynthesis. Success in this would potentially reveal novel enzymology, as well as providing fresh insight into the evolution of these sophisticated assembly-line pathways.

The initial questions to be addressed were whether [4+2] cyclase homologues are also present in the biosynthetic gene clusters of tetronasin and tetromadurin, and whether they are essential for their biosynthesis, given that no [4+2] cycloaddition is predicted to occur. Structural comparison of the polyether tetronate [4+2] cyclase homologues with those from spirotetronate pathways might help to shed light on their catalytic roles. It was hoped to use the putative cyclases to reconstitute (portions of) the post-PKS pathway *in vitro*, and hence to establish the timing and mechanism of cyclohexane and tetrahydropyran formation in polyether tetronate biosynthesis.

Chapter 2: Materials and methods

2.1 General methods

2.1.1 Chemical reagents

All chemicals and biomaterials were purchased from Sigma-Aldrich (USA) or Thermo Fisher (USA) unless otherwise stated. All chemicals were of analytical grade and all organic solvents were of HPLC quality. All water used in this study, unless stated otherwise, was distilled and purified though a Milli-Q water purification system (Millipore, USA). The purified tetronasin standard was a gift from Dr. Hui Hong.

2.1.2 Biological reagents.

All restriction endonucleases and DNA ligases were purchased from ThermoFisher Scientific (USA). T5 exonuclease and the Phusion™ and Q5™ high-fidelity DNA polymerases were purchased from New England Biolabs (USA). For standard PCR applications, such as colony PCR, BioMix Red™ was used (Bioline, USA). All enzymes were stored at −20 °C and used according to the manufacturer's instructions. Proteinase K and lysozyme were purchased from Sigma-Aldrich (USA).

2.1.3 Antibiotic solutions

Antibiotic solutions were made up to 1000x concentration stock solutions and filter-sterilised using a 0.22 μm filter before being stored at −20 °C. The final concentration of each antibiotic used is presented in **Table 2.1**. The antibiotics were then added to liquid medium or molten solid medium as required.

Antibiotic	Solvent	Final concentration used
Ampicillin	H_2O	100 µg/mL
Kanamycin	H_2O	50 µg/mL
Chloramphenicol	Ethanol	34 µg/mL
Apramycin	H_2O	50 µg/mL
Nalidixic acid	0.15 M NaOH	25 µg/mL

Table 2.1 Antibiotic concentrations used in this work

2.1.4 Oligonucleotides

All oligonucleotide primers were ordered from Sigma Aldrich (USA) in a lyophilised form. A master stock was created by dissolving each primer in TE buffer (10 mM TrisCl, 0.5 M EDTA, pH 8) to a final concentration of 100 µM. Working primer stocks were made to 10 µM in H_2O. All primers were stored at −20 °C.

Primer	Sequence (5'-3')
Primers used for creating pYH7 gene deletion constructs (using Gibson assembly)	
tsn11_Up_Fw	TGATCAAGGCGAATACTTCACGGCTCCGGCACCCCCG
tsn11_Up_Rv	CGTCGGCCGGTCCGGGTTCTTCCACGTGCCGC
tsn11_Dn_Fw	AGAACCCGGACCGGCCGACGCCGTCG
tsn11_Dn_Rv	CCGCGCGGTCGATCCCCGCATGGAGACGCAGGTTCCGGTGC
tsn15_Up_Fw	GGGACTGATCAAGGCGAATACTTCATATGTGGCGTGCTCGCTGAAGTAC
tsn15_Up_Rv	TGCGCGTTCGGTGGCGGGACGTTTCCTTCTGTCGGGCGG
tsn15_Dn_Fw	GAAACGTCCCGCCACCGAACGCGCACGGACGACGG
tsn15_Dn_Rv	GGGACCCGCGCGGTCGATCCCCGCATATGGGGTACGGCCGGGGGAGAAGC
tmn8_Up_Fw	GGGACTGATCAAGGCGAATACTTCAGCCGAGCTGCTGCTGGAACTGCTGC
tmn8_Up_Rv	GCGGCGCGCGACGCGGCGCCGACGCGGTCATGC
tmn8_Dn_Fw	GCGTCGGCGCCGCGTCGCGCGCCGCAGCC
tmn8_Dn_Rv	GGGACCCGCGCGGTCGATCCCCGCACTTGATCAGGCCGGTGATGCCGGCC
mad10_Up_Fw	GGGACTGATCAAGGCGAATACTTCAGCCGAGCTGCTGCTGGAACTGCTGC
mad10_Up_Rv	GCGGCGCGCGACGCGGCGCCGACGCGGTCATGC
mad10_Dn_Fw	GCGTCGGCGCCGCGTCGCGCGCCGCAGCC
mad10_Dn_Rv	GGGACCCGCGCGGTCGATCCCCGCACTTGATCAGGCCGGTGATGCCGGCC
mad31_Up_Fw	GGGACTGATCAAGGCGAATACTTCACCGGCAGCGGCGGAAGC
mad31_Up_Rv	CGACGATGTAGATCGGCGACACGCCGG
mad31_Dn_Fw	GGCGTGTCGCCGATCTACATCGTCGGAACGGG

mad31_Dn_Rv	GGGACCCGCGCGGTCGATCCCCGCACACCGGTGCCTCAGCG

Primers used for screening actinomycete gDNA for the targeted coding-frame deletions

tsn11_KO_Fw	ACCCAACTGACCATGCCC
tsn11_KO_Rv	GTGAACGTGGCGATGACC
tsn15_KO_Fw	ACTTCCAGCTCGAACTCCTG
tsn15_KO_Rv	TACGTTCAAGGAGTTCGGCT
tmn8_KO_Fw	GGTCTCAGCCCGTCGTTC
tmn8_KO_Rv	TGTGCTGCTCGACCTGTC
mad10_KO_Fw	CCTTCAAGAGATCCCCGAGG
mad10_KO_Rv	CGTCATCGGAACTCCCTGG
mad31_KO_Fw	CTCCTTCTGTTTGGGTGACG
mad31_KO_Rv	GAGTGCTTCAGGTGAGAACG

Primers used for creating pET28a(+) protein expression constructs (restriction sites underlined)

tsn11_NdeI_Fw	GGGGCATATGGAGATTCCTTTGACTGGCACCGTC
tsn11_XhoI_Rv	GGGGCTCGAGTCAGGCGGTGTCGCCGAAC
tsn15_NdeI_Fw	GGGGCATATGACCACTTCCATCGATCCGACG
tsn15_XhoI_Rv	GGGCTCGAGTCAGCGGCGCAGCTCGTACCG
ΔNtsn15_NdeI_Fw	TGGATTCCCATATGACCTACAACCCCGTCATCGA
tmn9_NdeI_Fw	GGGGCATATGAGTGAGCCGGTCGTCGT
tmn9_XhoI_Rv	GGGGCTCGAGTCAGCCGGTCGGGGC
tmn8_NdeI_Fw	GGGGCATATGCCTGCTCTCCTGAGCGGTGC
tmn8_XhoI_Rv	GGGGCTCGAGTCAGTCGCAGAGCGCAATGG
mad31_NdeI_Fw	GGGGCATATGGGCATGACCTCGCTGCGGCCTGG
mad31_XhoI_Rv	GGGGCTCGAGTCATCCGAAGGAAATGGAACTGAAGC
Fre_NdeI_Fw	GGGGGCATATGACAACCTTAAGCTGT
Fre_XhoI_Rv	GGGGCTCGAGTCAGATAAATGCAAACGC

Primers used for creating pIB139 genetic complementation constructs (using Gibson assembly)

pIB139_tsn11_Fw	CCACATATGTTGGGGATCCTATGGAGATTCCTTTGACTGGC
pIB139_tsn11_Rv	CGCGCGCGGCCGCGGATCCTTCAGGCGGTGTCGCCGAA
pIB139_tsn15_Fw	CCACATATGTTGGGGATCCTATGACCACTTCCATCGATCCG
pIB139_tsn15_Rv	CGCGCGCGGCCGCGGATCCTTCAGCGGCGCAGCTCG
pIB139_tmn8_Fw	CCACATATGTTGGGGATCCTGTGCCTGCTCTCCTGAGC
pIB139_tmn8_Rv	CGCGCGCGGCCGCGGATCCTTCAGTCGCAGAGCGCAAT
pIB139_tmn9_Fw	CCACATATGTTGGGGATCCTATGAGTGAGCCGGTCGTCG
pIB139_tmn9_Rv	CGCGCGCGGCCGCGGATCCTTCAGCCGGTCGGGGC
pIB139_mad10_Fw	AGGATCCACATATGTTGGGGATCCTATGAGCGATTCGGTTG
pIB139_mad10_Rv	GATATCGCGCGCGGCCGCGGATCCTTCACGCCGCAGGTTCC
pIB139_mad31_Fw	CCACATATGTTGGGGATCCTATGGGCATGACCTCGC
pIB139_mad31_Rv	CGCGCGCGGCCGCGGATCCTTCATCCGAAGGAAATGG

Primers used for creating site-specific point mutations

tsn15_E109A_Fw	TCCTGGTGACGGCGAAGATCTACAACTCC
tsn15_E109A_Rv	GGAGTTGTAGATCTTCGCCGTCACCAGGA
tsn15_D122A_Fw	GAGATCGGCGCGGCCACCGGGCGGCTGC
tsn15_D122A_Rv	GCAGCCGCCCGGTGGCCGCGCCGATCTC
tsn15_W190A_Fw	GCGGCACCCTGTCCGCGGTCATCAACTCC
tsn15_W190A_Rv	GGAGTTGATGACCGCGGACAGGGTGCCGC
tsn11_D282A_Fw	CGTCTTCGTCGCCGGCGCAGCCGCACACCTGCACG
tsn11_D282A_Rv	CGTGCAGGTGTGCGGCTGCGCCGGCGACGAAGACG

2.1.5 Plasmids used

All plasmids were stored in H_2O and kept at −20 °C (**Table 2.2**).

Name	Description	Source
pYH7	tsr, bla, aac(3)IV, cos, oriT, PT7, PT3, oripIJ101, oriColEI. Used for coding-frame deletions in *Streptomyces* by homologous recombination	[111]
pET28a(+)	kan, PT7, oripBR322, lacI, *N*-terminal His6-tag, oriF1. Used for the overexpression of genes in *E. coli* BL21	Novagen
pIB139	aac(3)IV, oriT, attP (s in *N*-terminal His6-tag, orIntegrative Plasmid used for *in trans* genetic complementations.	[206]
pYH7-*tsn11*	Gene deletion construct for *tsn11*	This study
pYH7-*tsn15*	Gene deletion construct for *tsn15*	This study
pYH7-*tmn8*	Gene deletion construct for *tmn8*	This study
pYH7-*mad10*	Gene deletion construct for *tsn15*	This study
pYH7-*mad31*	Gene deletion construct for *mad31*	This study
pIB139-*tsn11*	Construct for the heterologous expression of *tsn11*	This study
pIB139-*tsn15*	Construct for the heterologous expression of *tsn15*	This study
pIB139-*tmn8*	Construct for the heterologous expression of *tmn8*	This study
pIB139-*tmn9*	Construct for the heterologous expression of *tmn9*	This study
pIB139-*mad10*	Construct for the heterologous expression of *mad10*	This study
pIB139-*mad31*	Construct for the heterologous expression of *mad31*	This study
pET28a(+)-*tsn11*	*E. coli* overexpression construct for *tsn11*	This study
pET28a(+)-*tsn15*	*E. coli* overexpression construct for *tsn15*	This study
pET28a(+)-*tmn8*	*E. coli* overexpression construct for *tmn8*	This study
pET28a(+)-*tmn9*	*E. coli* overexpression construct for *tmn9*	This study
pET28a(+)-ΔN-*tsn15*	*E. coli* overexpression construct for ΔN-*tsn15*. A version of tsn15 lacking its first 11 amino acids.	This study
pET28a(+)-*tsn15*-Q84A	*E. coli* overexpression construct for *tsn15*-Q84A	GenBank
pET28a(+)-*tsn15*-R89A	*E. coli* overexpression construct for *tsn15*-R89A	GenBank
pET28a(+)-*tsn15*-Y91F	*E. coli* overexpression construct for *tsn15*-Y91F	GenBank
pET28a(+)-*tsn15*-E109A	*E. coli* overexpression construct for *tsn15*-E109A	This study
pET28a(+)-*tsn15*-D122A	*E. coli* overexpression construct for *tsn15*-D122A	This study
pET28a(+)-*tsn15*-T140A	*E. coli* overexpression construct for *tsn15*-T140A	GenBank
pET28a(+)-*tsn15*-S126A	*E. coli* overexpression construct for *tsn15*-S126A	GenBank
pET28a(+)-*tsn15*-N151A	*E. coli* overexpression construct for *tsn15*-N151A	GenBank
pET28a(+)-*tsn15*-Q164A	*E. coli* overexpression construct for *tsn15*-Q164A	GenBank
pET28a(+)-*tsn15*-W190A	*E. coli* overexpression construct for *tsn15*-W190A	This study
pET28a(+)-*tsn15*-Y202F	*E. coli* overexpression construct for *tsn15*-Y202F	GenBank

Table 2.2 Plasmids used in this study.

2.1.5 Bacteria strains used

All bacteria were grown on sterilised liquid or solid medium (**Table 2.3** and **Table 2.4**). A laminar flow hood and sterile technique was used when working with all strains.

Name	Characteristics	Source
E. coli BL21 (erist	F– ompT gal dcm lon hsd SB (rB- mB) λ(DE3 [lacI lacUV5-T7 gene 1 ind1 sam7 nin5]). *E. coli* strain used for heterologous protein expression from pET-28a(+).	Novagen (USA)
E. coli ET12657/pUZ8002	(F- dam -13::Tn9 dcm -6 hsd M hsd R recF143 zjj -202::Tn10 gal K2 gal T22 ara 14 pac Y1 xyl -5 leu B6 thi -1); pUZ8002: tra, neo, RP4. Donor strain for conjugation between *E. coli* and *actinomycetes*.	[207]
E. coli NovaBlue	*endA1 hsdR17* (r_{K12}^- m_{K12}^+) *supE44 thi-1 recA1 gyrA96 relA1 lac* F'[*proA*$^+$*B*$^+$*lacI*^{q}ZΔ*M15*::Tn*10*] (Tet^R). *E. coli* strain used for general DNA manipulation and cloning.	Novagen (USA)

Table 2.3 *E. coli* strains used in this study.

Name	Characteristics	Source
Actinomadura verrucosispora	Tetromadurin producer	PF[168]
Streptomyces sp. NRRL 11266	Tetronomycin producer	PF[166]
S. longisporoflavus 83E6	Tetronasin producer	PFL
S. longisporoflavus Δtsn11	*S. longisporoflavus* 83E6 in which 1356 nt of the *tsn11* coding frame has been deleted. Producer of T-16.	This study
S. longisporoflavus Δtsn15	*S. longisporoflavus* 83E6 in which the entire 621 bp *tsn15* coding frame has been deleted	This study
S. longisporoflavus Δtsn11 Δtsn15	*S. longisporoflavus* 83E6 in which 1356 nt of the *tsn11* coding frame and the entire 621 bp *tsn15* coding frame were deleted. Producer of T-16.	This study
S. longisporoflavus Δtsn11-*tsn11*	*S. longisporoflavus* Δtsn11 containing the integrative pIB139-*tsn11* plasmid. Produces tetronasin.	This study
S. longisporoflavus Δtsn15-*tsn15*	*S. longisporoflavus* Δtsn11 containing the integrative pIB139-*tsn15* plasmid.	This study
S. longisporoflavus Δtsn11-*tmn9*	*S. longisporoflavus* Δtsn11 containing the integrative pIB139-*tmn9* plasmid.	This study
S. longisporoflavus Δtsn11-*mad10*	*S. longisporoflavus* Δtsn11 containing the integrative pIB139-*mad10* plasmid.	This study
S. longisporoflavus Δtsn15-*tmn8*	*S. longisporoflavus* Δtsn15 containing the integrative pIB139-*tmn8* plasmid.	This study
S. longisporoflavus Δtsn15-*mad31*	*S. longisporoflavus* Δtsn15 containing the integrative pIB139-*mad31* plasmid.	This study
S. sp NRRL 11266 Δtmn8	*S.* sp NRRL 11266 83E6 in which the entire 573 bp *tmn8* coding frame has been deleted	This study
A.verrucosispora Δmad10	*A.verrucosispora* in which 906/1404 bp of the mad10 coding frame has been deleted. Producer of M-17.	This study
A.verrucosispora Δmad31	*A.verrucosispora* in which 291/561 bp of the mad10 coding frame has been deleted	This study

Table 2.4 Actinomycetes used in this study.

2.1.7 Bacteria culture media

All bacteria culture media was sterilised by autoclaving at 121°C, 1.5 bar, for 20 min. Tryptone, yeast extract, and agar were all purchased from Formedium (UK). Soy flour was purchased from Genesee Scientific (USA). D-mannitol was purchased from Alfa Aesar (USA). Dextrin was purchased from Acros Organics (USA). Tryptic soy broth was purchased from BD (USA). pH was adjusted using HCl and NaOH, measured using a Mettler Toledo 340 pH meter (USA).

LB medium
10 g/L tryptone
5 g/L yeast extract
10 g/L NaCl
H_2O to final volume (pH 7.0)

For solid LB medium 15 g/L agar was added before autoclaving.

2TY medium
16 g/L tryptone
10 g/L yeast extract
5 g/L NaCl
H_2O to final volume (pH 7.0)

SOC medium
20 g/L tryptone
5 g/L yeast extract
2.5 mM KCl
10 mM NaCl
10 mM $MgCl_2$**
20 mM D-glucose**
H_2O to final volume (pH 7.0)

**Added post-autoclave

M9 minimal medium

0.2% w/v D-glucose

2 mM $MgSO_4$

0.1 mM $CaCl_2$

100 µg/mL thiamine

X µg/mL antibiotic solution*

1x M9 salts**

20 mM NH_4Cl

0.2x trace metals***

20 µM $FeCl_3$****

H_2O to final volume (pH 7.0)

The M9 salts were autoclaved with 0.78x volume of H_2O. The remaining components were added post-autoclave.

*added according to Table 2.1

**5x M9 salts recipe

250 mM Na_2HPO_4

110 mM KH_2PO_4

50 mM NaCl

H_2O to final volume

Sterilized by autoclaving

***1000x Trace metal solution recipe

20 mM $CaCl_2$

10 mM $MnCl_2$

10 mM $ZnSO_4$

2 mM $CoCl_2$

2 mM $CuCl_2$

2 mM $NiSO_4$

2 mM Na_2MoO_4

2 mM H_3BO_3

Dissolved in H_2O and filter-sterilised.

**** The $FeCl_3$ solution was filter sterilised and made fresh for each batch of M9 medium.

SFM

20 g/L soy flour*

20 g/L D-mannitol

20 g/L agar

Tap water to final volume

*The soy flour was autoclaved first in 300 mL of tap water, after which it was filtered through a cloth to remove large insoluble components. The D-mannitol and agar was added to the filtrate, brought up to 1 L using tap water, and autoclaved again.

TSBY

30 g/L tryptic soy broth

103 g/L sucrose

5 g/L yeast extract

H_2O to final volume

Tap water medium (TWM)

5 g/L D-glucose

10 g/L sucrose

5 g/L tryptone

2.5 g/L yeast extract

0.036 g/L EDTA

15 g/L agar

Tap water to final volume (pH 7.1)

Oatmeal agar

20 g/L oatmeal*

H_2O to final volume

*The oatmeal was autoclaved first in 300 mL of H_2O after which it was filtered through a cloth to remove large insoluble components. H_2O was added to bring the volume up to 1 L and autoclaved again.

Tsn-medium A[208,209]

30 g/L tryptic soy broth

3 g/L $CaCO_3$

D4 trace elements (added post autoclave)*

H_2O to final volume

Tsn-medium B

30 g/L tryptic soy broth

3 g/L $CaCO_3$

20 g/L agar

100 g/L dextrin

H_2O to final volume

D4 trace elements*

*D4 trace elements recipe:

4 mg/L $FeSO_4$

4 mg/L $ZnSO_4$

0.6 mg/L $CuSO_4$

0.4 mg/L $MnSO_4$

0.4 mg/L $KMoO_4$

H_2O to final volume

2.2 Microbiological methods

2.2.1 Growth and maintenance of *E. coli* strains

E. coli liquid cultures were grown in LB, 2TY, or M9 medium at 37 °C, 200 rpm in conical flasks. *E. coli* solid cultures were grown at 37 °C on LB agar. Antibiotics were added as appropriate. For the long term storage of *E. coli* cultures, an equal volume of 50% glycerol was added to an aliquot and stored at −80 °C.

2.2.2 Growth and maintenance of actinomycete strains

All actinomycete mycelial liquid cultures were grown at 30 °C 200 rpm in conical flasks. The liquid medium was filled to no more than 1/5 of the total flask volume. A steel spring was placed at the bottom of each flask that extended around the inside face, preventing the culture from clumping. Foam springs were used to seal the flask. Actinomycete solid

cultures were grown at 30 °C for 4-7 days. For the long-term storage of actinomycete mycelial cultures, an equal volume of 50% glycerol was added and stored at −80 °C. Spores of *A. verrucosispora* were stored in 20% glycerol and stored at −80 °C.

2.2.3 Preparation of chemically competent *E. coli* cells

Transforming Buffer I (pH 5.8)	Transforming Buffer II (pH 7.0)
30 mM potassium acetate	10 mM NaMOPS (pH 7.0)
50 mM $MnCl_2$	75 mM $CaCl_2$
10 mM $CaCl_2$	KCl 10 mM
15 % glycerol (v/v)	15 % glycerol (v/v)

E. coli cultures were inoculated in 3 mL of LB + appropriate antibiotics and grown overnight at 37 °C 200 rpm. The following day 100 µL of the overnight culture was used to inoculate 50 mL of LB containing 10 mM $MgCl_2$ and any appropriate antibiotics. The cells were grown to an A_{600} of 0.3-0.4, after which they were placed on ice for 20 min before being spun at 2700 x *g*, 2 °C for 10 min. The supernatant was decanted and the cells were resuspended in 0.8x volumes of ice-cold transforming buffer I then placed on ice for 2 h. The cells were then spun down again at 2700 x *g*, for 10 minutes at 2 °C before being resuspended in 0.1x volumes of ice-cold transforming buffer II. Aliquots of the cells were then snap-frozen and stored at −80 °C.

2.2.4 Transformation of chemically competent cells

A frozen aliquot of chemically competent cells was defrosted on ice, and 50-100 ng of plasmid DNA was added. The cells were kept on ice for 30 min before heat-shocking at 42 °C for 60 sec, after which 250 µL of SOC medium was added. The cells were incubated at 37 °C 220 rpm for 1 h before being plated onto LB agar containing any appropriate antibiotics and incubated overnight at 37 °C.

2.2.5 Conjugation of actinomycete strains

The DNA methylation-deficient ET12567 (pUZ8002) cells were transformed with the plasmid to be transferred into the actinomycete host and grown to an A_{600} of 0.4-0.6 in 2TY media containing selective antibiotics at half the usual concentration (see **Table 1**). The ET12567 cells were then spun down and washed twice in 20 mL of 2TY media before being resuspended in 300 µL of 2TY. The actinomycete strain to be conjugated (either as mycelia or spores) was prepared following the protocol from *Practical Streptomyces Genetics*[210]. After gentle mixing, the two bacteria types were plated onto 35 mL of SFM containing 10 mM $MgCl_2$ and left at 30 °C for 12-20 h. The surface of the plate was then flooded with 1 mL of MQ water containing 35 µL of apramycin (50 mg/mL) and 25 µL nalidixic acid (25 mg/mL). The plate was then incubated at 30 °C to promote the growth of exconjugants. Exconjugants were verified by restreaking onto SFM containing 30-50 µg/mL apramycin and 25 µg/mL nalidixic acid.

2.3 Molecular biology methods

2.3.1 Extraction of genomic DNA from actinomycetes

For the preparation of genomic DNA (gDNA), 250 µL of actinomycete mycelia was centrifuged in a 1.5 mL microcentrifuge tube to pellet the cells and the supernatant was discarded. The cell pellet was resuspended in 500 µL of SET buffer (20 mM TrisCl, 75 mM NaCl, 75 mM EDTA, pH 7.2) containing 10 µL of lysozyme solution (50 mg/mL). After incubation at 37 °C for 1 h, 60 µL of 10% SDS (w/v) and 10 µL of proteinase K (20 mg/mL) were added and the tube was incubated at 55 °C for an additional 2 h. The sample was then mixed with 300 µL of 5 M NaCl and 500 µL of chloroform and centrifuged at 2200 x *g* for 15 min. The aqueous upper layer was transferred to a fresh microcentrifuge tube using a T1000 pipette tip with the end cut off to avoid shearing the gDNA. To precipitate the gDNA, 0.6x volumes of isopropanol was added followed by gentle mixing. The precipitated gDNA was washed twice in 70% ethanol before being air dried for 5-10 min. The gDNA pellet was dissolved in 100-200 µL of distilled water.

2.3.2 Isolation of plasmid DNA

The extraction of plasmid DNA from *E. coli* was performed using a Thermo Fisher GeneJET Plasmid Miniprep Kit™ (Thermo Fisher, USA) following the manufacturer's instructions.

2.3.3 Quantification of DNA and protein concentration

DNA was quantified by measuring the A260 nm absorbance using a NanoDrop ND-1000 Spectrophotometer (Thermo Scientific, USA). DNA quality was assessed by measuring the 260/280 nm and 260/230 nm ratios. Protein concentration was calculated by measuring its A280 nm absorbance and its extinction coefficient (calculated using web.expasy.org/protparam/). For the quantification of proteins bound to an FAD-cofactor the Bradford reagent (Biorad, USA) was used according to the manufacturer's instructions. A bovine serum albumin (BSA) standard curve (presented in **Appendix Figure 1**) was prepared to use with the Bradford reagent.

2.3.4 Restriction enzyme digestion of DNA

DNA (both linear and circular) was digested using ThermoFisher Scientific™ restriction enzymes (USA). Restriction digestion mixtures were set up according to the manufacturer's instructions and incubated at 37 °C for 1–3 hours. Following digestion restriction enzymes were inactivated by heating at 65 °C for 20 min or using a DNA clean and concentrator kit (Zymo Research, USA).

2.3.5 Agarose gel-electrophoresis

50x TAE buffer
242 g/L Tris base
5.7 % (v/v) acetic acid
6 g/L EDTA

The size separation of DNA fragments was performed using agarose gel-electrophoresis. A 1% (w/v) agarose solution (100 mL) was prepared in 1x TAE buffer. Midori Green™ DNA stain (Bulldog Bio) was added to the molten gel according to the manufacturer's

instructions. The gel was poured into a gel-caster and a well-comb inserted. Gels were run at 120 V, 400 mA, for 30-40 min. GeneRuler 1kb Plus DNA ladder (Thermo Scientific) was used as a DNA size standard.

2.3.6 Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE)

Solutions required:

Resolving gel recipe (single gel)
1.1 mL ddH_2O
2.5 mL 30% acrylamide mix
1.3 mL 1.5 M TrisCl pH 8.8
0.05 mL 10% (w/v) Sodium dodecyl sulfate
0.05 mL 10% ammonium persulfate (made fresh)
0.0002 mL Tetramethylethylenediamine

Stacking gel recipe (single gel)
0.68 mL ddH_2O
0.17 mL 30% acrylamide mix
0.13 mL 1.5 M TrisCl pH 8.8
0.01 mL 10% (w/v) Sodium dodecyl sulphate
0.01 mL 10% ammonium persulfate*
0.0001 mL Tetramethylethylenediamine (TEMED)

4x SDS Loading Buffer
200 mM TrisCl pH 6.8
400 mM DTT
8% SDS (w/v)
0.4% bromophenol blue (w/v)
40% glycerol (v/v)

10x SDS run buffer
0.114% glycine (w/v)
0.03% tris base (w/v)
0.01 % SDS (w/v)

The 15% resolving gel was made first and added into ~3/4 of the gel cast, the remainder of which was filled with isopropanol to ensure the surface of the gel set uniformly. The gel was then left to set for approximately 20 min at room temperature. Once set, the isopropanol was poured off and the 15% stacking gel loaded onto top of the resolving gel. A comb was inserted into the stacking gel before it was left to set for 20 min. Protein samples were loaded into the gel after being mixed with 1x loading buffer and heated at 98 °C for 10 min. The gels were immersed 1x SDS run buffer and electrophoresis was carried out at 250 V, 20 mA, for 40-50 min. Precision Plus Dual Colour ladder (Bio-Rad) was used to provide protein size markers. Protein bands were visualised using InstantBlue™ (Expedeon, UK).

2.3.7 Polymerase chain reaction

Polymerase chain reaction (PCR) amplifications of DNA were performed using a

GeneAmp™ 9700 PCR machine (PE Applied Biosystems). The two polymerase mixtures typically used were 2X Phusion™ High Fidelity Master Mix with GC Buffer (New England Biolabs) and 2X BioMix Red™ (Bioline). A typical PCR was carried out in 1X polymerase mastermix, 0.5 µM forward primer, 0.5 µM of each primer and 5% (v/v) DMSO, and made up to the final volume in H_2O. 2-10 ng of DNA was typically used for a PCR template. The PCR thermal parameters were taken from the Phusion™ and BioMix™ manufacturer's instructions with consideration for the melting point (T_m) of each primer used in the reaction (typically ~3 °C lower than the lowest primer T_m). For colony PCR, a small piece of *E. coli* or actinomycete colony was added directly to the PCR tube. PCR amplification mixtures were prepared on ice and added to the PCR thermocycler once it had reached its starting temperature. A common two-step PCR protocol for the amplification of high GC DNA fragments is presented in **Figure 2.1**.

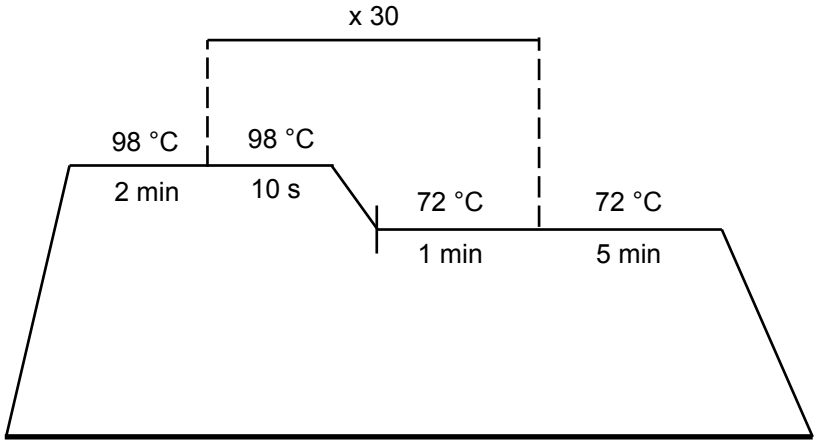


Figure 2.1 Typical PCR thermal parameters used for the amplification of actinomycete genomes.

2.3.8 Heterologous protein expression from *E. coli*

Proteins were heterologously expressed from pET-28a(+) within *E. coli* BL21 (λDE3) cells. A pET-28a(+) construct containing a target gene cloned in-frame to the *N*-terminal His_6 tag (MGSSHHHHHHSSGLVPRGSH) was transformed into chemically competent *E. coli* BL21 (λDE3). A single transformant was selected to inoculate a 3 mL overnight culture of LBK (LB + 50 µg/mL kanamycin) media. The following day the overnight culture was added to 400-1000 mL of LBK media and grown at 37 °C 200 rpm to an A_{600} of 0.5-0.6 before being

placed on ice for 20 min. Isopropyl β-D-1-thiogalactopyranoside (IPTG) was then added to a final concentration of 0.5 mM and the cells were incubated at 22 °C 200 rpm overnight. The cells were then spun down and frozen at −80 °C until needed.

2.3.9 Purification of *N*-terminal His_6-tagged proteins

Buffers required:

Protein Binding buffer
20 mM TrisCl pH 7.9
0.5 M NaCl
5 mM imidazole

Protein Elution Buffer
20 mM TrisCl pH 7.9
0.5 M NaCl
200-600 mM imidazole

Protein Wash buffer
20 mM TrisCl pH 7.9
0.5 M NaCl
100 mM imidazole

Protein Strip buffer
20 mM TrisCl pH 7.9
0.5 M NaCl
100 mM imidazole
100 mM

A frozen aliquot of *E. coli* BL21 (λDE3) cells in which protein had been induced as described (**Section 2.3.8**) was resuspended in 30-40 mL of protein binding buffer. A SigmaFast EDTA-free™ protease inhibitor tablet was dissolved in the resuspended cells. For small-scale protein purifications (<400 mL) cell lysis was performed using sonication. For large-scale purifications (>400 mL) lysis was performed using an EmulsiFlex-C5 (Avestin, Canada) according to the manufacturer's instructions. The lysed cells were centrifuged at 35,000 x g, 4 °C for 20 min. The supernatent was run through a column containing 1 mL of charged nickel-NTA resin (Qiagen, DE). After the supernatant had been applied to the column, non-specifically bound proteins were washed off the resin using 10 mL of protein binding buffer followed by 6 mL of protein wash buffer. The fusion-protein was eluted from the column using 5 mL of elution buffer. A preparative ÄKTA purifier (GE Healthcare Life Sciences) Superdex 200 column was used to purify the protein further, exchange the buffer into 20 mM TrisCl pH 7.9, 100 mM NaCl, and remove the imidazole. An Amicon Ultra-15 Centrifugal Filter Unit (Merck Millipore, USA) was used to concentrate

the protein following preparative gel-filtration. The nickel column was cleaned after each use with 5 mL of strip buffer.

2.3.10 Sanger DNA sequencing

All DNA cloning reactions (such as ligations and isothermal assemblies) were confirmed by Sanger DNA sequencing. All sequencing was performed by the DNA Sequencing Facility of the Department of Biochemistry, University of Cambridge, UK.

2.3.11 Gibson assembly and method of creation gene-deletion constructs

Gibson assembly[211] of two linear DNA fragments into a linearised vector was performed using Hot Fusion DNA assembly mix[212]. The pYH7 vector was linearised using the restriction enzyme *Nde*I. The general method used to create a gene-deletion construct was to PCR amplify 2 kb regions upstream and downstream of the gene of interest. The PCR primers used to amplify the 2 kb fragments contained regions of overlap with *Nde*I-linearised pYH7 and the other 2 kb fragment, enabling all three to be seamless joined using Hot Fusion DNA assembly[212]. Each 2 kb fragment was added in three-fold molar excess of the linearised pYH7 vector. Typically the total DNA used in each reaction was 100 ng. The assembled DNA product was then transformed into chemically competent *E coli* NovaBlue cells. Colony PCR was used to identify *E. coli* clones containing the correctly assembled insert, which were then fully sequenced using Sanger sequencing.

2.3.12 Site-directed mutagenesis

Site-directed mutagenesis was performed using the QuikChange II mutagenesis guidelines (Agilent, USA). The gene to be mutated was first cloned into the small vector pUC18. Complementary PCR primers each containing the desired mutation were designed and used to amplify the entire plasmid. The methylation-specific restriction enzyme D*pn*I was added to the amplified plasmid to digest the methylation template plasmid DNA. The remaining reaction was then transformed into chemically competent *E. coli* NovaBlue. Sanger sequencing was then used identify the clones containing the desired mutation. Only the Tsn15-E109A, Tsn15-D122A, and Tsn15-W190A were made using this method, the rest were ordered directly from GenScript (China).

2.3.13 Ligation of DNA

DNA ligase was purchased from New England Biolabs (USA) and used according to the manufacturer's instructions. Typical reactions contained 50 ng of digested vector with 3-6 fold molar excess of the insert. The ligation reaction was incubated at 16 °C for 1-4 hours before being transformed into chemically competent *E. coli* NovaBlue cells. Colony PCR was performed to identify the clones containing the correct ligation product. All ligation constructs were confirmed using Sanger sequencing.

2.3.14 Buffer exchange using PD-10 columns

PD-10 columns were used change the buffer of a purified protein/desalt a protein solution. PD-10 columns from GE Healthcare Life Sciences (USA) were used according to the manufacturer's instructions.

2.3.15 Analytical ultracentrifugation of proteins

Analytical ultracentrifugation (AUC) sedimentation was used to measure the molecular weight of a protein in solution. Purified protein was prepared as described in **Sections 2.3.8 and 2.3.9** and concentrated to 1-2 mg/mL in a final volume of 800 µL. AUC experiments were performed using an Optima XL-I (Beckman Coulter) centrifuge fitted with an An60 Ti four-hole rotor. Absorbance data was acquired in the continuous mode at time intervals of 170 s and at a rotor speed of 40,000 rpm, at 20 °C with systematic noise subtracted, but without averaging. The density and viscosity of the buffer (20 mM TrisCl pH 7.9, 100 mM/150 mM NaCl) and the partial specific volume of the protein were both calculated using Sednterp[213]. The multi-component sedimentation coefficient distributions were obtained from 128 scans by direct boundary modelling of the Lamm equation using Sedfit v.14.1. All AUC experiments were performed under the guidence of Dr. Katherine Stott (Biophysics Facility, Department of Biochemisty, University of Cambridge).

2.3.16 HPLC-MS analysis of proteins

For the analysis of small molecules, a HPLC (Hewlett Packard, Agilent Technologies 1200 series) coupled to a mass spectrometer (Thermo Finnigan MAT LTQ) was used. The HPLC was fitted with a 250 mm x 4.6 mm 5µm C4 reverse-phase column (5µ OSD3, 100Å. Phenomenex, USA). The mobile phase was aqueous 0.1% trifluoroacetic acid (TFA) and acetonitrile (0.1% TFA). The mobile phase flow rate was 0.3 mL/min with the following gradient: 0-20 min, 5-95% acetonitrile (0.1% TFA) 20-24 min, 95% acetonitrile (0.1% TFA), 25-26 min, 95-5% acetonitrile (0.1% TFA).

2.3.17 Protein crystallisation

Proteins to be crystallised were purified using Ni-affinity followed by ÄKTA size-exclusion chromatography. Next, a Mosquito™ (TTP Labtech) liquid handling robot was used to mix 0.2 µL of protein sample into 0.2 µL of the different pre-prepared crystallisation buffers. The crystallisation plates, each containing 96 different conditions, screened were:

- The Classics Suite (Qiagen, DE)
- JCSG *plus* (Molecular Dimensions, UK)
- Wizard Classic 1&2 (Emerald Biosystems, USA)
- Wizard Classic 3&4 (Emerald Biosystems, USA)
- PACT *premier*™ HT-96 (Molecular Dimensions, UK)
- pH Clear I (Qiagen, DE)
- The PEGs suite (Qiagen, DE)
- Morpheus™ (Molecular Dimensions, UK)

Droplets were photographed under white and UV light using a Rock Imager 1000 (Formulatrix, USA) to detect crystallisation events. The Tsn15 crystal selected for analysis was grown from 15 mg/mL protein in 0.1 M PCTP (0.1 M each sodium proprionate, sodium cacodylate trihydrate, and bis-Tris propane) buffer pH 6, 25% (w/v) PEG 1500. 26% (v/v) ethylene glycol was used as cryo-protectant and the crystal was flash frozen in liquid nitrogen.

To create a selenomethionine labelled version of Tsn15 (SeMet-Tsn15) *E. coli* BL21 (λDE3) cells transformed with pET28a(+)-*tsn15* were cultured in M9 medium. Once the culture had reached A_{600} = 0.5, 0.1 g/L L-lysine, 0.1 g/L L-threonine, 0.1 g/L L-phenylalanine, 0.05 g/L L-leucine, 0.05 g/L L-isoleucine, and 0.05 g/L L-valine were added to the growing culture, followed by 0.06 g/L of L-selenomethionine. SeMet-Tsn15 was purified and exchanged into buffer containing 20 mM Tris-Cl pH 7.9, 100 mM NaCl, 2 mM EDTA, and 2 mM tris(2-carboxyethyl)phosphine (TCEP), then crystallised as before. Crystals were harvested from the crystallisation drop and transferred into a cryoprotecting solution containing the crystallisation solution and 26% ethylene glycol.

The following was performed by Fernanda Paiva (University of São Paulo, Brazil). The Tsn15-ligand complex was crystallised using hanging-drop, vapor-diffusion method at 18 ºC. To obtain the protein complex, a solution composed of Tsn15 protein (15 mg/mL in 20 mM Tris-Cl pH 7.9 buffer, 0.5 M NaCl) and 10 mM of T-22 was incubated for 10 minutes and then crystallised in 0.1 M PCTP pH 6.0 buffer, 27% PEG 1500.

2.3.18 Circular dichroism

Circular dichroism (CD) spectra were collected using an Aviv Biomedical Model 410 (USA) circular dichroism spectrometer. The CD spectrometer was operated by the Department of Biochemistry (University of Cambridge) Biophysics facility. The buffer of each purified protein was exchanged to 10 mM KH_2PO_4, pH 7.9, as TrisCl-based buffers strongly absorb in the far UV. The concentration of each protein used was between 0.1 - 0.2 mg/mL. After the CD spectra were collected the data was normalised to the number of peptides present in each sample (mean residue ellipticity Θ): calculated using:

$$\Theta_{mre}(\lambda)\ (\text{deg.cm}^2/\text{dmol}) = \frac{\text{millidegrees}}{(\text{amino acids} - 1) \cdot \text{concentration (M)} \cdot \text{cell pathlength (cm)} \cdot 10}$$

2.4 Analysis of polyether tetronates and related metabolites

2.4.1 HPLC-MS analysis of small molecules

For the analysis of small molecules, an HPLC (Hewlett Packard, Agilent Technologies 1200 series) coupled to a mass spectrometer (Thermo Finnigan MAT LTQ) was used. The HPLC-MS was fitted with a 250 mm x 4.6 mm 5µm C18 reverse-phase column (5µ OSD3, 100Å. Phenomenex, USA). The mobile phase comprised of 20 mM ammonium acetate and methanol. The mobile phase flow rate was 0.7 mL/min with the following gradient: 0-5 min, 5-75% methanol; 5-30 min, 75-95% methanol, 30-34 min, 95% methanol 35-36 min, 95-5% methanol. Normalised collision energy of 35% was used for the molecular fragmentation of tetronasin.

2.4.2 Small-scale extraction of polyether tetronates

To check the production of polyether tetronate metabolites from liquid culture, a 500 µL aliquot was removed and extracted using 500 µL of ethyl acetate. The ethyl acetate layer was then removed and evaporated under reduced pressure to yield a dry organic extract. The dried organic extract was dissolved in 1 mL of methanol and centrifuged at 20,000 g for 20 min, before HPLC-MS analysis. For the extraction of polyether metabolites produced by cultures grown on solid medium, the agar was cut into tiny cubes and extracted by submerging in ethyl acetate. The ethyl acetate was evaporated and the organic extract was redissolved 1 mL of methanol and centrifuged at 20,000 g for 20 min to remove any particulates, before HPLC-MS analysis.

2.4.3 NMR data collection

All NMR data was collected by the Department of Chemistry (University of Cambridge) service team using a 500 MHz NMR spectrometer (Bruker, USA).

2.4.4 Isolation of T-16 for NMR analysis

T-16 was isolated from *S. longisporoflavus* Δtsn11. Well-grown *S. longisporoflavus* Δtsn11 colonies were inoculated into 5 mL of tsn-medium-A in a 25 mL conical flask and grown for two days at 30 °C and 200 rpm. The cultures were then inoculated into flasks containing 50 mL of tsn-medium-A and grown for an additional two days at 30 °C 200 rpm or until a thick mycelial culture had formed. These cells were plated on 100 12 x 12 cm agar plates each containing 50 mL of tsn-medium-B in a 250 mL conical flask and grown for seven days at 30 °C. The agar was then cut into small squares and combined in a large glass flask where it was extracted three times by submerging in ethyl acetate (1.5 L each time). The ethyl acetate was evaporated under reduced pressure to yield 7 g of a brown crude organic extract. The crude extract was dissolved in 8 mL of methanol and loaded onto eight 10g/70 mL C18 reverse-phase Isolute cartridges (Biotage, Sweden) according to the manufacturer's instructions. Fractions were eluted from the columns using a solution of 20 mM ammonium acetate, 65% methanol. Fractions containing T-16 were identified by HPLC-MS and pooled together (1.15 g in total) before a second round of purification on another C18 cartridge. The final purification was performed using an Infinity II semipreparative HPLC (Agilent, USA) fitted with a Phenomenex (USA) C18 Prodigy column (5 µm ODS-3 100Å, 250 x 10 mm). Gradient elution of T-16 was achieved using a mobile phase of 5 mM ammonium acetate and methanol with a flow rate of 3 mL/min: 0-5 min, 5-75% methanol; 5-15 min, 75-85% methanol; 15-19 min, 85-100% methanol; 19-20 min, 100-5% methanol. Fractions containing T-16 were identified by detecting its characteristic chromophore (λ_{max}=236) nm and pooled together, followed by solvent evaporation under reduced pressure. The dried extract was dissolved in ethyl acetate to remove the ammonium acetate, followed by freeze-drying to yield purified ca. 12.5 mg T-16. For structural determination, T-16 was dissolved in deuterated chloroform and analysed using a 500 MHz DCH Cryoprobe Spectrometer (Bruker, USA).

2.4.5 Isolation of T-22 for NMR analysis

T-22 was isolated from an *in vitro* incubation containing T-16 and purified Tsn11. To obtain sufficient T-16 for the assay, 300 more 12 x 12 cm tsn-medium-B agar plates containing *S. longisporoflavus* Δtsn11 were grown. The assay mixture (60 mL in total) contained: 20 mM Tris-Cl, 100 mM NaCl, pH 7.9 containing 50 µM Tsn11, 400 µM of T-16, and 5% v/v

methanol. After incubation at 30 °C for 1.5 h, the reaction mixture was extracted six times using 30 mL of ethyl acetate. The ethyl acetate was evaporated to yield a dried organic extract that was redissolved in methanol. The extract was chromatographed on semipreparative HPLC described above. Fractions containing intermediate T-22 were identified by detecting its characteristic chromophore (λ_{max} = 236 nm) and pooled together for solvent evaporation. The dried residue was dissolved in ethyl acetate to remove the ammonium acetate, followed by freeze-drying to yield purified T-22 (3.3 mg). For structural determination, T-22 was dissolved in deuterated methanol and analysed using a 500 MHz DCH NMR Cryoprobe Spectrometer (Bruker, USA).

2.4.6 *in vitro* enzyme activity assays

In vitro activity assays using Tsn11 and Tsn15 (or their homologues) were performed in 20 mM TrisCl buffer, 100 mM NaCl, 5% v/v methanol pH 7.9. Typical, reactions were 100 µL and contained 200 µM of T-16 or T-22. Tsn11 and/or Tsn15 were added to a final concentration of 5 µM. The reactions were carried out at 30 °C for 1 h unless stated otherwise. The reactions were terminated by adding 400 µL of methanol (to denature the enzymes) before being completely dried under reduced pressure. The dried extract was redissolved in 100 µL methanol and analysed by an Infinity II semipreparative HPLC (Agilent, USA) fitted with a Phenomenex (USA) C18 Prodigy column (5 µm ODS-3 100 Å, 250 x 10 mm) using a gradient program of 20 mM ammonium acetate and increasing methanol at a flow rate of 3 mL/min: 0-5 min, 5-75% methanol; 5-15 min, 75-85% methanol; 15-19 min, 85-100% methanol; 19-20 min, 100-5% methanol.

For analysis of the Tsn11 sodium dithionite assays and the Tsn15 point mutation assays, a Poroshell 120, EC-C18, 27 µM, 46 x 100 mm (Agilent, USA) column eluted with 5 mM ammonium acetate and methanol at 1 mL/min: 0-5 min, 5-75% methanol; 5-15 min 75-95% methanol; 15-19 min, 95% methanol; 19-20 min, 95-5% methanol. When the Tsn11 activity assay was conducted in the presence of sodium dithionite, 100 µM of T-16 was used with 1 µM of Tsn11 and 1 mM of sodium dithionite. Parallel reactions were carried out and terminated at 0 min, 2 min, 4 min, and 8 min, respectively. For all *in vitro* assays, the identity of intermediates T-16 and T-22 were confirmed by their unique UV chromophores and MS^3 fragmentation patterns.

2.4.7 High-resolution mass spectrometry (HRMS)

All ESI high-resolution mass (HRMS) spectrometry was performed by the Department of Chemistry, University of Cambridge mass spectrometry service.

2.4.8 Hydrogen-deuterium exchange

To exchange the hydroxyl protons of a small molecule with deuteriums, the small molecule was dissolved in a 50:50 mix of deuterated methanol and D_2O. The mixture was then directly injected into the mass spectrometer.

2.4.9 Interception and detection of PKS-bound polyketide intermediates

Method written and performed by Dr Manuela Tosin and Robert Jenkins, University of Warwick, United Kingdom. All *S. longisporoflavus* strains were grown in duplicate 10 mL tsn-medium A for 2 days at 30 °C at 200 rpm. These were incubated at 30 °C for 5 days. After the first day of incubation, the probe (methyl 6-decanamido-2-fluoro-3-oxohexanoate, final concentration: 1 mM) was dissolved in 80 µl of MeOH and added portionwise over 4 days (20 µl each day, days 2-5). Control liquid cultures omitting methyl 6-decanamido-2-fluoro-3-oxohexanoate were also prepared (in duplicate copy). After 5 days of fermentation, the liquid cultures were extracted with ethyl acetate (20 mL). The extracts were concentrated and the residues were redissolved in HPLC-grade methanol (1 mL) for mass spectrometry analysis. HPLC-HR-ESI-MS analyses of *S. longisporoflavus* extracts were carried out on an LTQ-T Orbitrap Fusion instrument. Reverse phase chromatography was used to separate the mixtures prior to MS analysis. Two columns were utilised: an Acclaim PepMap µ-precolumn cartridge 300 µm i.d. x 5 mm 5 µm 100 Å and an Acclaim PepMap RSLC 75 µm x 15 cm 2 µm 100 Å (Thermo Scientific, USA). The columns were installed on an Ultimate 3000 RSLCnano system (Dionex, USA). Mobile phase buffer A was 0.1% aqueous formic acid and mobile phase B was composed of 100% acetonitrile containing 0.1% formic acid. Samples were loaded onto the µ-precolumn equilibrated in 50% aqueous acetonitrile containing 0.1% trifluoroacetic acid for 8 min at 10 µL min^{-1} after which compounds were eluted onto the analytical column following a 75 min gradient for which the mobile phase B concentration was increased from 50% B to 80% over 15 min,

then maintained at 80% B for 50 minutes, then decreased to 50% over 1 min, followed by a 9 min wash at 50% B. Species were analysed by electrospray ionisation mass spectrometry. Survey scans of precursors from 150 to 1500 *m*/*z* were performed at 60K resolution (at 200 *m*/*z*) with a 4×10^5 ion count target. Tandem MS was performed by isolation at 1.6 Th with the quadrupole, HCD fragmentation with normalised collision energy of 32, and rapid scan MS analysis in the ion trap. The MS/MS ion count target was set to $2x10^5$ and the maximum injection time was 50 ms. A filter targeted inclusion mass list was used to select the precursor ions.

2.5 Bioinformatics and software

2.5.1 Solving protein crystal structures

X-ray diffraction data for SeMet-Tsn15 were collected at the Diamond Light Source (DLS) I04 beamline, Oxford, UK. The X-ray data processing was performed using XDS[214] and scaled using Aimless[215] (both from the CCP4 suite[216]). The phases for SeMet-Tsn15 was solved using AutoSol[217]. The structural model was initially built by AutoBuild[218]. The models were refined using Phenix.refine[219] and COOT[220]. The stereochemical quality of the models was assessed using MolProbity[221]. The final protein structure was visualised using PyMOL (The PyMOL Molecular Graphics System, Version 2.0, Schröinger, LLC).

The structure of the Tsn15-ligand complex was solved by collaborators Feranda Paiva and Dr. Marcio Dias. X-ray data was collected for the Tsn15-ligand complex was collected at PETRAIII, beamline 13, Hamburg, Germany. The phases for the Tsn15-ligand complex was determined by molecular replacement using the program Phaser[222] from Phenix suite[223]. The SeMet-Tsn15 structure was used as the search model for molecular replacement. The protocol for model building and refinement was the Tsn15-ligand complex was the same as for SeMet-Tsn15.

2.5.2 Construction of phylogenetic trees

Maximum-likelihood (based on the Le-Cascuel model[224]) phylogenetic trees were created using MEGA7[225] by following the protocol described in (226). 1000 bootstrap replicates were performed. All positions with less that 95% site-coverage were not included.

2.5.3 Protein software

Protein sequence alignments were performed using ClustalOmega[227].
Protein extinction coefficients and molecular weights were estimated using ProtParam[228]

BLAST was used for all query-based sequence alignments[229].

The PyMOL Molecular Graphics System (Schrödinger, LLC) was used to visualise protein crystal structures.

2.5.4 DNA software

All *in silico* DNA cloning experiments and interpretation of sequence chromatogram data was performed using SnapGene (GSL Biotech).

Where appropriate, PCR primers were designed with the assistance of Primer3[230]. The melting properties of primers was assessed using OligoAnalyzer[231]

2.5.5 Chemistry software

Chemical structures and molecular weights were calculated using ChemDraw Professional 16.0 (Perkin-Elmer Informatics). Mass spectra were visualised using Xcalibur (Thermo Fisher Scientific). Deconvolution of protein mass spectra was performed using ProMass (Novatia, LLC).

Chapter 3: [4+2] cyclase homologues in polyether tetronate biosynthetic gene clusters

3.1 The tetromadurin (SF2487/A80577) biosynthetic gene cluster

The unexpected presence of the [4+2] cyclase homologue *tmn8* in the *tmn* cluster[157] prompted the search for *tmn8* homologues in the BGCs of tetronasin (*tsn*) and tetromadurin (*mad*). While the presence of a [4+2] cyclase homologue in only the *tmn* cluster could indicate that it is an unused, vestigial gene, presence in all three polyether tetronate BGCs would support a conserved and essential biosynthetic role. Given that the arrangement of the cyclohexane and tetrahydropyran rings is identical in tetronomycin, tetronasin, and tetromadurin it is reasonable to believe they form by a common mechanism. As a first step, it was necessary to identify the biosynthetic gene cluster for tetromadurin (the *mad* cluster) (**Figure 3.1**).

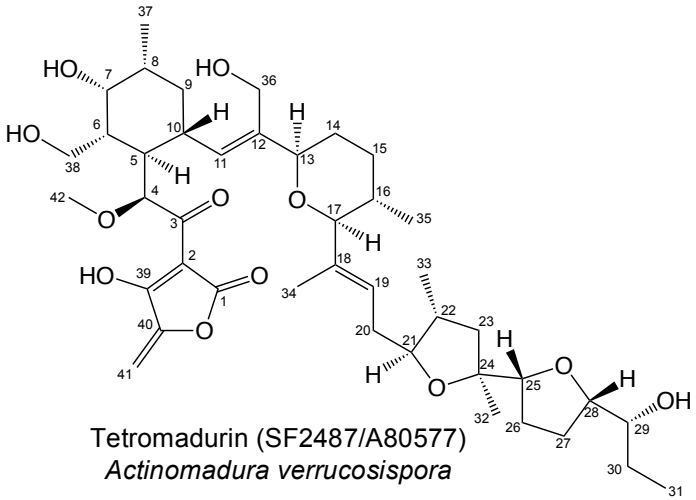


Figure 3.1 Structure of tetromadurin (SF2487/A80577) with atom numbering.

Tetromadurin is a type I polyketide polyether tetronate isolated independently from *Actinomadura verrucosispora* by two industrial research groups, under the designations SF2487[168] and A80577[169]. Tetromadurin contains similar elements to both tetronasin and tetronomycin. The stereochemistry of the cyclohexane and tetrahydropyran rings is identical to that of tetronasin, while the tetronate ring is tetronomycin-like, containing an

exocyclic double bond. In regards to bioactivity, tetromadurin is active against a range of gram-positive organisms[168] and inhibits HIV-1 replication in H9 cells[232]. A 2005 patent described the antimalarial activity of tetromadurin, although no further developments have occurred since[233].

A strain of the tetromadurin producer *A. verrucosispora* NRRL B-18236 was obtained from the Agricultural Research Service Culture Collection (NRRL). Whole-genome sequencing of the strain was performed by the Department of Biochemistry DNA sequencing facility (University of Cambridge) using combined shotgun and long-range mate pair MiSeq data. The genome was assembled into a single scaffold (10.215 Mbp) by Dr. Markiyan Samborskyy and analysed by AntiSMASH 4.0[234] to identify biosynthetic gene clusters. A summary of the results from the AntiSMASH 4.0 analysis is presented in **Figure 3.2**.

Actinomadura verrucosispora NRRL B-18236	
Genome size	10.215 Mbp
No. of scaffolds	1
No. of annotated BGCs	26
Type I PKS	3
NRPS	3
NRPS-PKS	4
Terpene	6
RiPPs	6
Bacteriocin	3
Type III PKS	1
Non-NRPS siderophore	1

Figure 3.2 Predicted biosynthetic gene clusters in *A. verrucosispora.* AntiSMASH 4.0[234] was used to identify the biosynthetic gene clusters in the genome of *A. verrucosispora*. Dr. Markiyan Samborskyy.performed the AntiSMASH 4.0[234] analysis.

Three biosynthetic gene clusters containing type I PKSs were identified in the *A. verrucosispora* genome. The carbon backbone of tetromadurin was predicted to be composed from six acetate units, eight propionate units, one glyceryl-ACP unit, and an unusual (2*R*)-methoxymalonyl-ACP unit (**Figure 3.3**). Only one of the type I PKS BGCs in *A. verrucosispora*, 110 kbp in size and containing 32 genes, could feasibly produce tetromadurin (**Figure 3.4**). The exact boundaries of the *mad* cluster were decided on the basis of the presence of genes whose annotation suggested no obvious connection with tetromadurin biosynthesis.

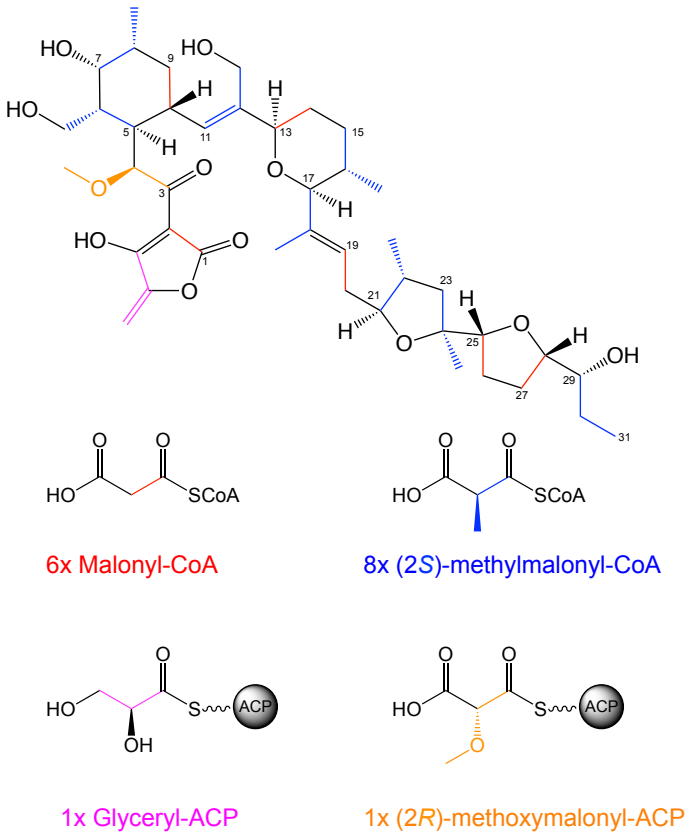


Figure 3.3 Proposed biosynthetic precursors of the tetromadurin carbon backbone. The carbon backbone of tetromadurin appears to be derived from six malonyl-CoA derived acetate units (red), eight (2*S*)-methylmalonyl-CoA derived propionate units (blue), one glyceryl-ACP (magenta), and one (2*R*)-methoxymalonyl-ACP (orange).

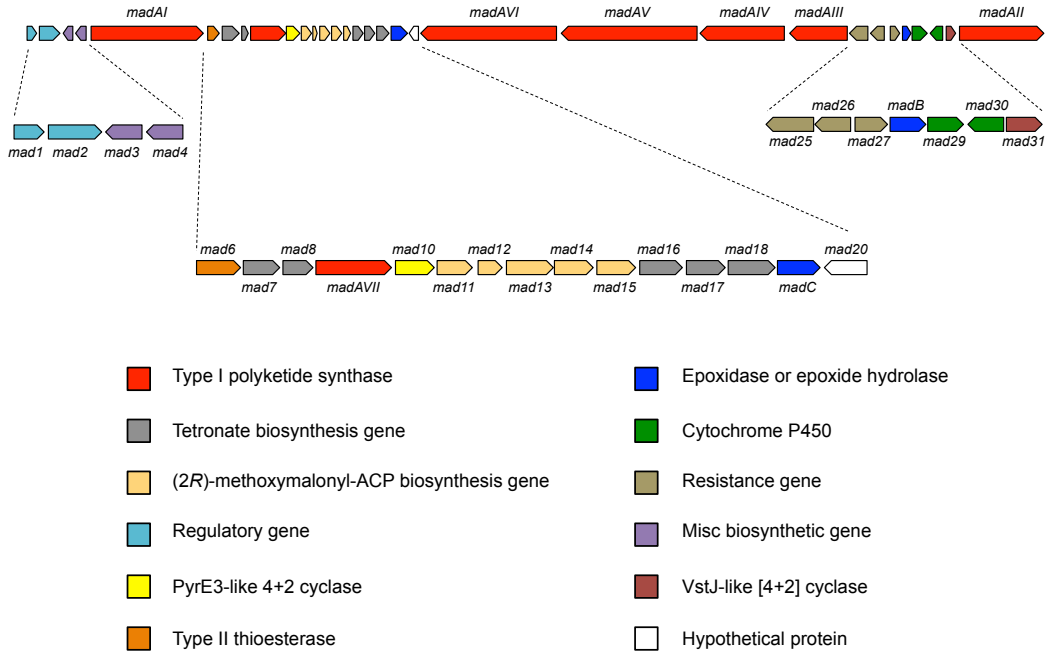


Figure 3.4 Biosynthetic gene cluster of the polyether tetronate tetromadurin (SF2487). The total size of the *mad* cluster is 110 kbp (genes not drawn to scale). A total of 32 were predicted to comprise the *mad* cluster. Genes are colour coded according to their predicted function.

Putative functions for the genes in the *mad* cluster were predicted using BLASTX analysis[229] and are listed in **Table 3.1**. Each gene type will be discussed in reference to its proposed role in tetromadurin biosynthesis.

Table 3.1 Predicted functions of genes in the tetromadurin (*mad*) biosynthetic gene cluster. ATa = malonyl-CoA selective AT domains. ATp = *2S*-methylmalonyl-CoA selective AT domains.

orf	Size*	Homologue and origin	Identity/ similarity (%)	Proposed function
mad1	130	WP_067469187, *Actinomadura macra*	79/90	SARP activator protein
mad2	937	KX263301 (MonH) *Streptomyces cinnamonensis* strain ST021	50/59	LuxR transcriptional activator
mad3	192	WP_017621714 (EntD), *Nocardiopsis gilve*	66/73	4'-phosphopantetheinyl transferase
mad4	285	WP_113699285 (CpdA), *Nonomuraea* sp. NEAU-YG30	74/80	Metallophosphoesterase
madAI	4869	AJE80656, *Streptomyces albus*	54/65	Type I PKS: KS^Q, ATp, ACP, KS, ATa, DH, KR, ACP, KS, ATa, DH, ER, ACP
mad6	256	WP_110701439 (GrsT) *Streptosporangium* sp. caverna	59/72	Type II thioesterase
mad7	642	WP_067456393, *Actinomadura macra*	66/73	FkbH-like glyceryl-S-ACP synthase
mad8	75	WP_031047730, *Amycolatopsis albispora*	63/80	Acyl-carrier protein
madAVII	1588	WP_093461893, *Streptomyces melanosporofaciens*	50/60	Type I PKS: KS, ATa, KR, ACP
mad10	467	ACR50781 (Tsn11), *Streptomyces longisporoflavus*	48/60	FAD-dependent monooxygenase/ putative [4+2] cyclase
mad11	288	WP_037917628 (FabB), *Streptomyces* sp. PCS3-D2	69/75	Glyceryl-S-ACP dehydrogenase
mad12	78	PPS67085, *Streptomyces* sp. 46	55/69	Acyl-carrier protein
mad13	362	WP_027945620 (CaiA), *Amycolatopsis taiwanensis*	71/78	2-hydroxy-3-oxopropionyl-S-ACP dehydrogenase
mad14	375	WP_055537123, *Streptomyces neyagawaensis*	71/80	FkbH-like glyceryl-S-ACP synthase
mad15	226	WP_086736314, *Streptomyces glaucescens*	64/78	SAM-dependent methyltransferase
mad16	342	WP_113696535, *Amycolatopsis albispora*	64/76	FabH 3-oxoacyl-ACP synthase III family protein
mad17	281	WP_055549000 (AceF), *Streptomyces kanamyceticus*	58/75	Acyltransferase
mad18	383	WP_067456402, *Actinomadura macra*	54/70	Dehydratase
madC	472	CCD31908, *Streptomyces albus*	55/70	Epoxidase
mad20	273	WP_119927171 *Streptosporangiaceae bacterium* YIM 75507	44/61	Hypothetical protein
madAVI	5623	WP_049717538, *Streptomyces caatingaensis*	50/62	Type I PKS: KS, ATp, DH, ER, KR, ACP, KS, ATp, DH, KR, ACP, KS, ATmethoxy, DH, KR, ACP
madAV	5670	WP_045867303, *Streptomyces* sp. NBRC 110027	51/63	Type I PKS: KS, ATp, DH, ER, KR, ACP, KS, ATp, DH, KR, ACP, KS, ATp, DH, KR, ACP
madAIV	3343	WP_104651095, *Streptomyces cinnamoneus*	54/66	Type I PKS: KS, ATp, DH, KR, ACP, KS, ATp, KR, ACP
madAIII	1819	WP_078876216, *Streptomyces* sp. 769	55/67	Type I PKS: KS, ATa, DH, KR, ACP
mad25	535	WP_092886900, *Actinopolymorpha cephalotaxi*	57/72	ABC transporter permease
mad26	312	WP_026416314, *Actinomadura oligospora*	69/81	ABC transporter ATP-binding protein
mad27	180	WP_092886896, *Actinopolymorpha cephalotaxi*	52/69	MarR family transcriptional regulator
madB	132	ACR50776 (TsnB), *Streptomyces longisporoflavus*	53/69	Epoxide hydrolase
mad29	386	ACR50783 (Tsn12), *Streptomyces longisporoflavus*	52/65	Cytochrome P450
mad30	400	WP_055419038, *Streptomyces pactum*	59/72	Cytochrome P450
mad31	186	BAF73716 (Tmn8), *Streptomyces* sp. NRRL 11266	29/45	Putative [4+2] cyclase
madAII	3978	WP_106436315, *Streptomyces bingchenggensis*	57/68	Type I PKS: KS, ATp, DH, KR, ACP, KS, ATp, DH, ER, ACP

* Amino acid residues

3.1.1 Polyketide synthase genes in the tetromadurin biosynthetic gene cluster

The *mad* cluster contains seven large genes (*madAI-madAVII*) encoding type I polyketide synthase multienzymes. Together, the seven enzymes collectively consist of one loading and 14 extension modules. No thioesterase domain is present in any of the modules, as expected for a tetronate pathway. The loading PKS module (MadAI_KS_LM) could be differentiated from extension modules by its KS^Q domain (**Figure 3.5**).

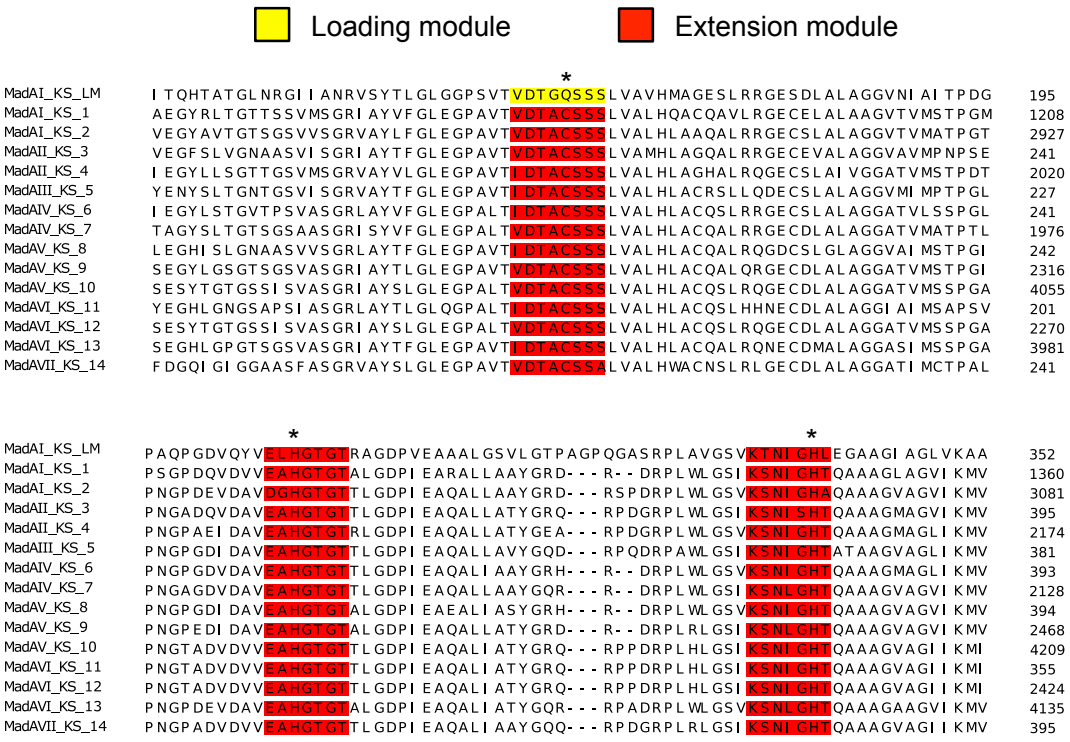


Figure 3.5 Alignment of the KS domains from the tetromadurin biosynthetic gene cluster. The KS domain of the loading module (MadAI_AT_LM) and each of the 14 extension modules (KS1-14) from the tetromadurin biosynthetic gene cluster were aligned using ClustalOmega[227]. The amino acids of the C-H-H catalytic triad are highlighted with an asterisk (*). The KS domain of the loading module contains a glutamine in place of the catalytic cysteine (highlighted in yellow).

The order in which the remaining six PKS multienzymes act was predicted by a detailed examination of their constituent extension modules to determine: (1) the predicted substrate selectivity of the AT domain; (2) the tally of reducing domains (KR, DH, and ER) present; and (3) whether the KR/ER domains would establish the correct stereochemistry of the α-methyl and/or β-hydroxyl groups. Only one arrangement of the PKS enzymes was predicted to produce a linear backbone related to tetromadurin, and the PKS multienzymes were labelled accordingly MadAI-MadAVII (**Figure 3.6**). With two exceptions, the predicted product of MadAI-MadAVII is identical to the linear intermediate predicted from the tetromadurin structure itself, as discussed below.

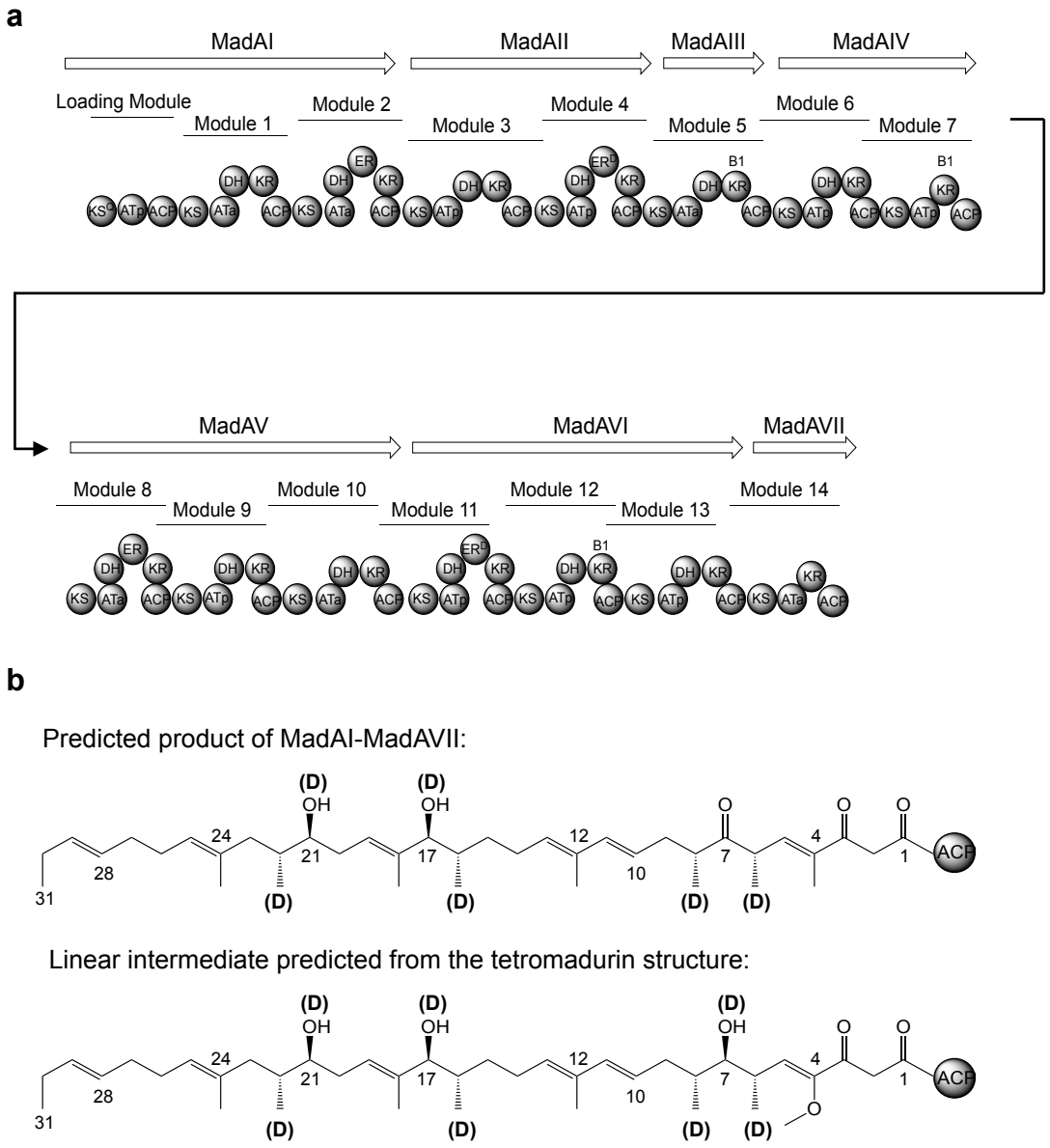


Figure 3.6 Hypothetical linear tetromadurin backbone. a, The proposed order of the seven PKS enzymes in the *mad* cluster to produce the linear tetromadurin backbone. ATa domains contain the amino acid motifs associated with malonyl-CoA incorporation. ATp domains contain the amino acid motifs associated with (2*S*)-methylmalonyl-CoA incorporation. ER^D domains are predicted to create a D-configured α-methyl group. B1 KRs are predicted to create D-configured α-methyl and β-hydroxyl groups. The KS domain in MadAI is a KS^Q domain. **b,** Top: the hypothetical linear tetromadurin intermediate predicted to be produced by the MadAI-MadAVII arrangement. Bottom: the hypothetical linear tetromadurin intermediate predicted from the modules of MadAI-MadAVII.

3.1.2 AT domains of the tetromadurin polyketide synthase enzymes

The substrate selectivity of the AT domains was predicted from the presence of specific amino acid motifs[80,91,92], particularly the presence of a (H/T/V/Y)AFH (acetate-incorporating), or (Y/V/W)ASH (propionate-incorporating) motif. The AT domain of the

loading module contains the hybrid HASH motif, previously found in modules with a relaxed selectivity that can incorporate both malonyl-CoA and (2*S*)-methylmalonyl-CoA[235]. However, no congener of tetromadurin with one fewer methyl group has been reported, indicating that (2*S*)-methylmalonyl-CoA is still specifically recruited by the loading module. The predicted AT selectivity of the 14 extension modules is also essentially in accordance with the structure of tetromadurin (**Figure 3.7**). The AT domain of module 13 (MadAVI_AT_13) contains a YASH motif, indicating selectivity for a (2*S*)-methylmalonyl-CoA. However, based on the structure of tetromadurin this module should actually incorporate (2*R*)-methoxymalonyl-ACP to form the C4 methoxy group. The selectivity motifs for AT domains that incorporate (2*R*)-methoxymalonyl-ACP are poorly defined, and these AT domains typically resemble either a malonyl-CoA or (2*S*)-methylmalonyl-CoA selective domain[236]. Despite this, it is likely that MadAVI_AT_13 really does incorporate (2*R*)-methoxymalonyl-ACP, given that the *mad* cluster contains the genes necessary to produce this unusual extension unit (**Section 3.1.5**).

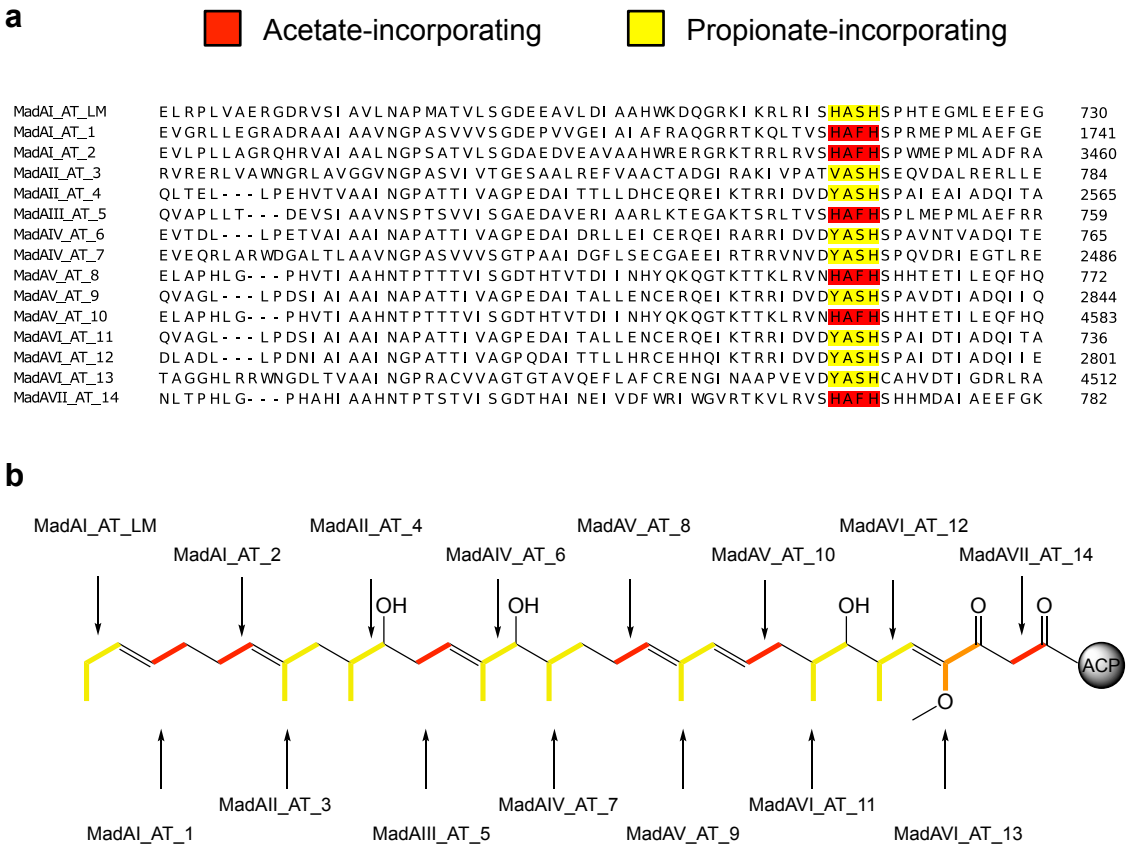


Figure 3.7 Alignment of the AT domains from the tetromadurin biosynthetic gene cluster. a, The AT domain of the loading module (MadAI_AT_LM) and each of the 14 extension modules (AT_1-14) from the *mad* cluster were aligned using ClustalOmega[227]. The amino acid motif that predicts the selectivity of the AT domain is highlighted in either red (acetate-incorporating) or yellow (propionate-incorporating). **b,** The hypothetical linear tetromadurin intermediate predicted by retrobiosynthetic analysis. Propionate units derived from (2*R*)-methoxymalonyl-ACP are coloured in yellow. Acetate units derived from malonyl-CoA are coloured in red. The methoxy group predicted to arise from (2*R*)-methoxymalonyl-ACP is coloured in orange.

3.1.3 KR, DH, and ER domains of the tetromadurin polyketide synthase enzymes

With one exception, the tally of reducing domains (KR, DH, and ER) present in each module is also consistent with the structure of tetromadurin. Modules 2, 4, 8, and 11 all contain a KR, DH, and ER domain, consistent with complete saturation at C9, C15, C23, and C27, respectively (**Figure 3.1**). Since modules 4 and 11 incorporate propionate units, the ER domains in these modules determine the configuration of the α-methyl groups at C8 and C22, respectively. These ER domains both contain amino acid motifs associated with the D-configuration (lacking the key tyrosine)[105], consistent with the D-configured α-methyl substituents expected at these positions (**Figure 3.8**).

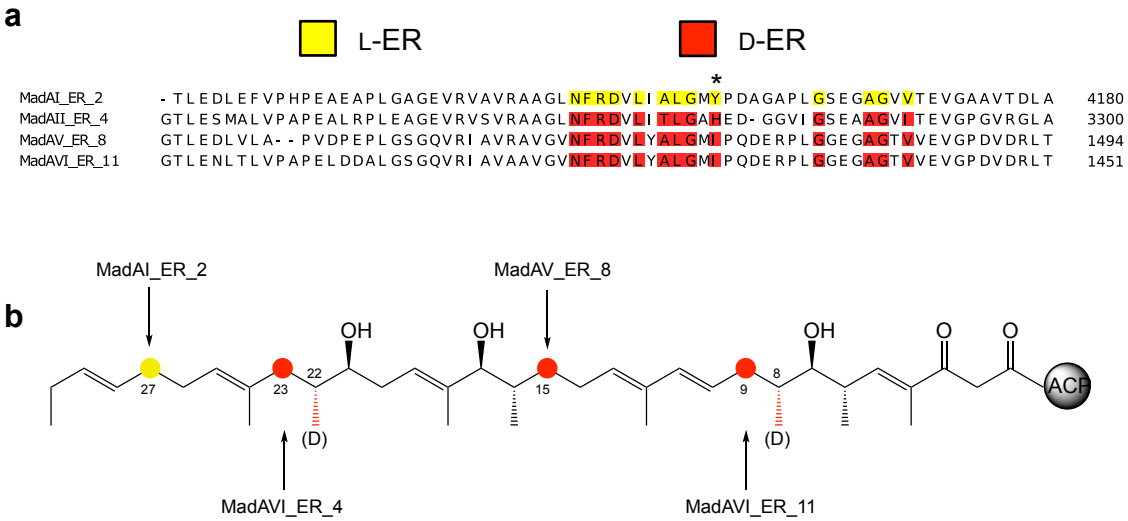


Figure 3.8 Alignment of the ER domains from the tetromadurin polyketide synthase enzymes. a, The ER domains from the tetromadurin biosynthetic gene cluster were aligned using ClustalOmega[227]. The tyrosine residue associated with the L-configuration at the α-position in the product is marked with an asterisk (*). The ER domains are highlighted either yellow (producing L-configuration at the α position) or red (producing D-configuration at the α position). Only MadAII_ER_4 and MadAII_ER_11 affect the final stereochemistry of tetromadurin, as their respective modules incorporate a propionate unit. **b,** The linear tetromadurin intermediate predicted from its final structure. The fully saturated atoms (C9, C15, C23, and C27) produced by an ER domain are highlighted either in yellow (L-configuration-determining ER) or red (D-configuration-determining ER).

The DH domains of modules 1, 3, 6, 9, 10, and 13 are all appropriately placed to form the double bonds at C29, C25, C19, C13, C11, and C5, respectively. Although module 5 contains a DH domain, C21 contains a hydroxyl group in the final tetromadurin structure, suggesting that MadAIII_DH_5 is inactive. Examination of the active sites of MadAIII_DH_5 confirmed it likely inactivity, as it is missing the tyrosine from the YGP motif. The tyrosine side chain of the YGP motif in DH domains is proposed to assist binding to the β-hydroxyl group of the substrate[65] and has been shown experimentally to be essential[237]. The remaining DH domains all appear to be active (**Figure 3.9**).

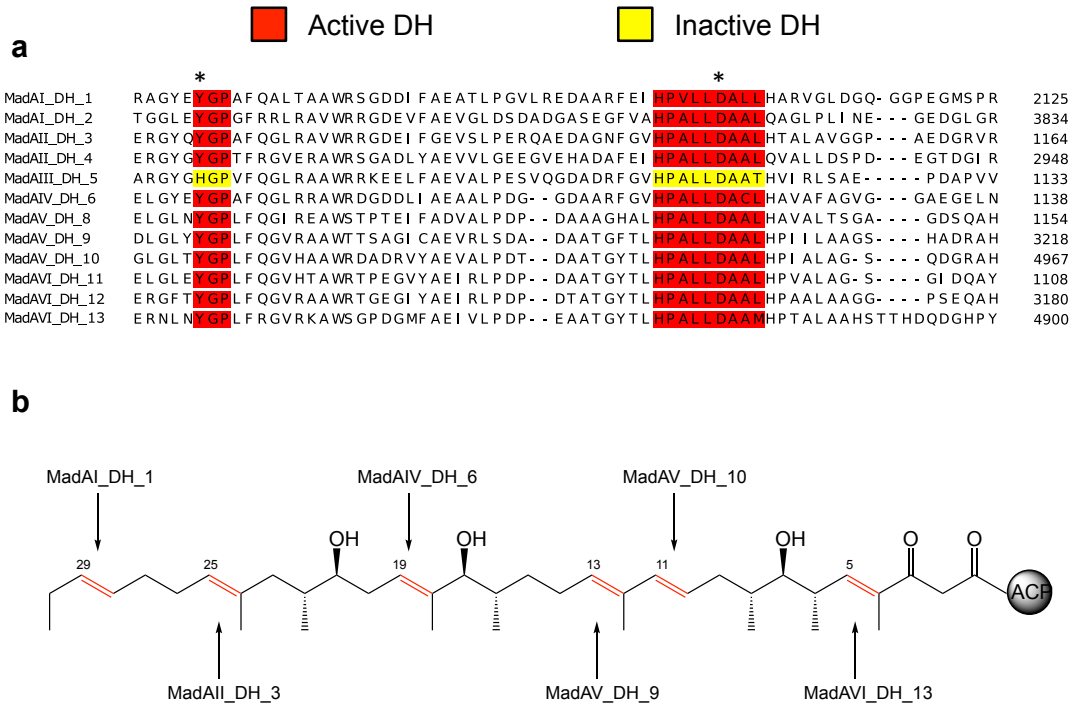


Figure 3.9 Alignment of the DH domains from the tetromadurin polyketide synthase enzymes. a, The DH domains of the extension modules 1-13 from the *mad* cluster were aligned using ClustalOmega[227]. The tyrosine of the **Y**GP motif and the catalytic aspartic acid in the HPALL**D**AAL motif are marked with an asterisk (*). DH domains predicted to be active are highlighted in red and those predicted to be inactive are highlighted in yellow. **b,** The linear tetromadurin intermediate predicted from its final structure. The double bonds produced by DH domains are highlighted in red.

Ketoredutase domains exert significant stereochemical influence over growing polyketide chains, determining the stereochemistry of the β-hydroxyl group and (where appropriate) the α-group of each extension unit. In the proposed MadAI-MadAVII module order, the KR domains of module 5, 7, and 12 are appropriately placed to form the hydroxyl groups at C7, C17, and C21. In the case of tetromadurin, all backbone β-hydroxyl groups and the α-methyl groups adjacent to these hydroxyls (C33, C34, and C36) are in the D-configuration, indicating the activity of B1 type KR domain[101]. The distinguishing features of a B1 KR domain are an (L,V,I)DD motif and the absence of a proline (a feature of A type KRs) two residues C-terminal of the catalytic tyrosine in A type KRs[101,238]. Protein sequence analysis confirmed that the KR domains from modules 5 and 7 are indeed B1, but surprisingly module 12 appears to be inactive (**Figure 3.10**).

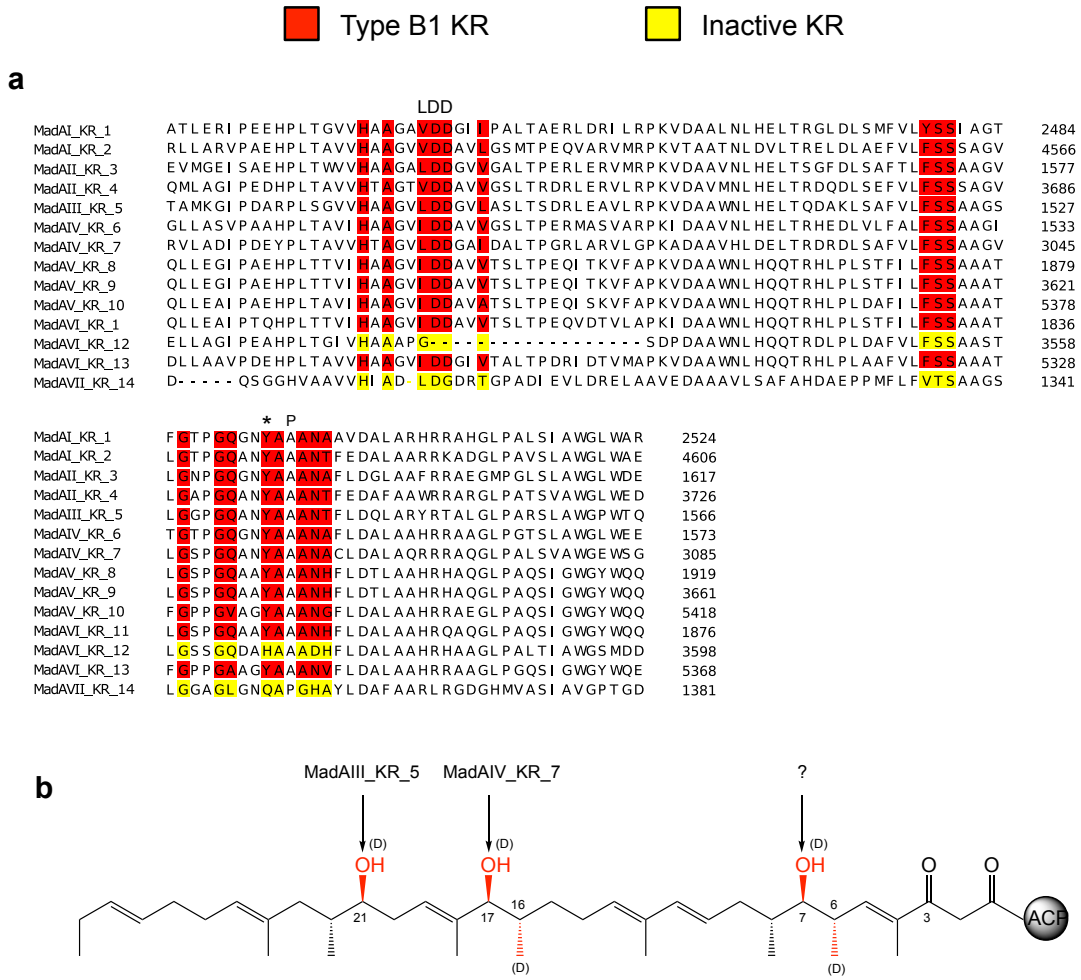


Figure 3.10 Alignment of the KR domains from the tetromadurin polyketide synthase enzymes. a, The KR domains of the extension modules 1-14 from the *mad* cluster were aligned using ClustalOmega [227]. The key features of a B1 type KR domain are highlighted: the (L/V/I)DD motif within the NADPH binding site (LDD), the catalytic tyrosine (*), and the locus of the proline found in A type KRs (P). All of the KR domains except KR_12 and KR_14 were predicted to be type B1. KR_12 and KR_14 both lacked the catalytic tyrosine residue, indicating they are inactive. KR_12 also had a significant deletion in its NADPH binding site. **b,** The linear tetromadurin intermediate predicted from its final structure. The three D-configured hydroxyls (C7, C17, and C21) predicted to form from a B1 type KR domain are highlighted in red. The D configuration of the C6 and C16 methyl groups is also predicted to be governed by B1 type KRs. The KR domain in the position corresponding to the C7 hydroxyl (MadAVI_KR_12) appears to be inactive, so the origin of this hydroxyl is unknown.

In fact, the KR domains of module 12 and 14 both appear to be inactive, lacking the key catalytic tyrosine[65,239], with module 12 also containing a large deletion in its NADPH binding site. While the inactivity of KR_14 is consistent with the C3 keto group of tetromadurin, the inactivity of KR_12 is not. An inactive KR domain at this position would result in a keto group at C7 (as shown in **Figure 3.6**) rather than the C7 D-hydroxyl group observed in the tetromadurin. To rule out this apparent inactivity being due to an error in genome sequencing/assembly, the region was specifically amplified using PCR and sequenced

again. Resequencing confirmed the original sequence was correct. Furthermore, despite possessing this mutated KR domain, the strain was still a competent producer of tetromadurin (see **Chapter 4**). How the C7 hydroxyl group forms in tetromadurin is therefore a mystery. Additionally confounding is that module 12 also contains a seemingly active DH domain that should replace the C7 D-hydroxyl by a C6-C7 *trans* double bond. Aside from this, the PKS modules present were an excellent fit for producing the carbon backbone of tetromadurin.

3.1.4 Genes for tetronate biosynthesis

The *mad* cluster contains homologues of all five glycerate-utilisation operon genes: *mad7*, *mad8*, *mad16*, *mad17*, and *mad18*[109,122,123]. *mad7* and *mad8* respectively encode a FkbH-like protein and standalone acyl carrier protein found in the biosynthetic pathways of other tetronates[111,122]. Mad7 likely catalyses formation of glyceryl-ACP from 1,3-bisphosphoglycerate using Mad8 as the ACP scaffold. Mad16 is the FabH-like protein known for catalysing tetronate formation and chain release[122]. Finally, Mad17 and Mad18 are homologues of the acyltransferase Agg4 and the dehydratase Agg5, respectively, responsible for exocyclic double bond formation in agglomerin biosynthesis[123]. Mad17 likely catalyses the acetylation of the C41 hydroxyl, followed by Mad17 catalysing elimination of the acyl group to form the exocyclic C41-C41 double bond (**Figure 3.11**).

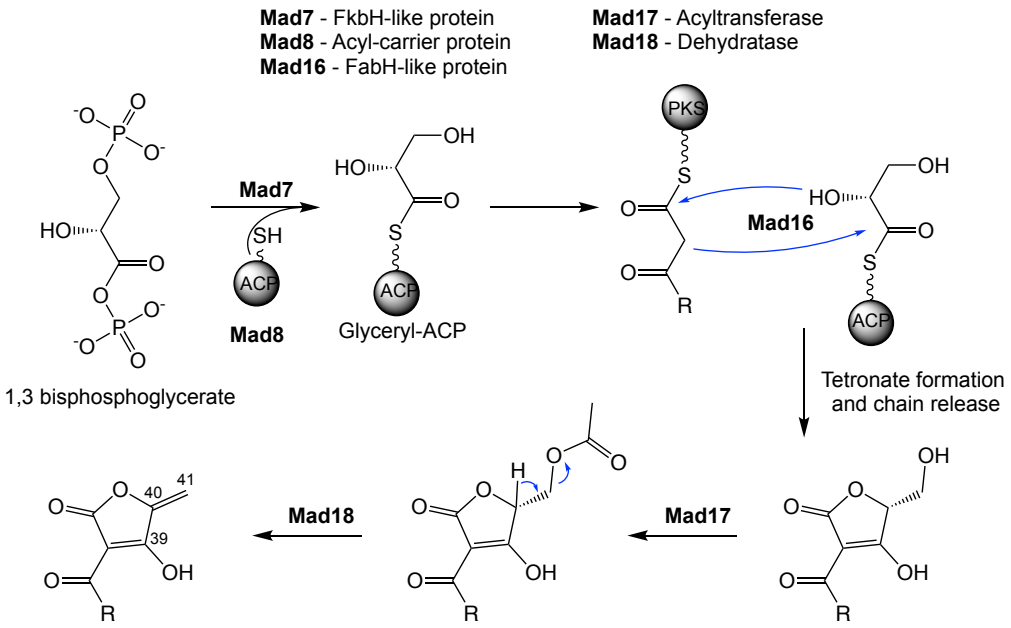


Figure 3.11 Proposed biosynthesis of the tetromadurin tetronate ring. The *mad* gene cluster contains the five genes of the glycerate-utilisation operon responsible for tetronate ring biosynthesis. The FkbH-like Mad7 catalyses formation of glyceryl-S-ACP using 1,3-bisphosphoglycerate and the ACP, Mad8. The FabH-like protein Mad16 then catalyses C-C and C-O bond formation to form the tetronate ring and release the polyketide chain from the PKS. Finally, the C40-C41 exocyclic double bond is formed *via* acetylation of the C41 hydroxyl by acetyltransferase Mad17, followed by elimination catalysed by the dehydratase Mad18. R = linear tetromadurin intermediate.

3.1.5 Genes for (2*R*)-methoxymalonyl-ACP biosynthesis

Several polyketide natural products are known to incorporate the unusual methoxymalonate extender unit derived from (2*R*)-methoxymalonyl-ACP[61]. The first example of this was the macrocyclic immunosuppressant FK520[236]. Five genes in the FK520 BGC were identified as responsible for the biosynthesis of (2*R*)-methoxymalonyl-ACP[236] from 1,3-bisphosphoglycerate. Homologues of these five genes have since been found in the BGC of other natural products known to incorporate (2*R*)-methoxymalonyl-ACP, such as the apoptolidins[240], concanamycin A[241], and geldanamycin[242,243]. The *mad* cluster also contains homologues of the same five genes: *mad11*, *mad12*, *mad13*, *mad14*, and *mad15* (**Figure 3.12**), which appear to be part of a single operon. Mad14 encodes a second FkbH-like protein, distinct from Mad8, that likely catalyses formation of a second pool of glyceryl-ACP using 1,3 bisphosphoglycerate and Mad12, another standalone acyl carrier protein. It is unknown if the two pools of glyceryl-ACP can interchangeably be used in tetronate formation and (2*R*)-methoxymalonyl-ACP biosynthesis. The primary hydroxyl

group of glyceryl-ACP undergoes two oxidation events to form (2*R*)-hydroxymalonyl-ACP. Based on their similarity to the dehydrogenase enzymes in the FK520 pathway, Mad11 is proposed to catalyse the first oxidation, forming 2-hydroxy-3-oxopropionyl-ACP. Mad13 then catalyses the second oxidation to form (2*R*)-hydroxymalonyl-ACP. The *O*-methyltransferase Mad15 likely converts (2*R*)-hydroxymalonyl-ACP into (2*R*)-methoxymalonyl-ACP. The exact timing of *O*-methylation is uncertain, and it may precede the oxidation steps[244].

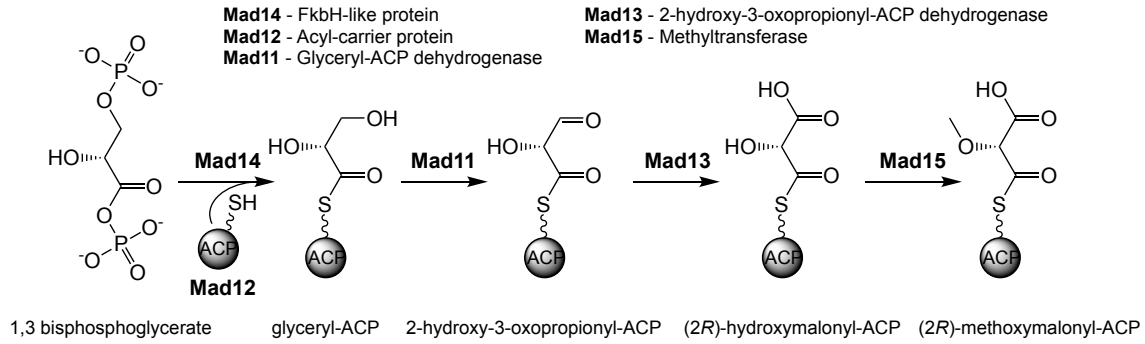


Figure 3.12 Proposed biosynthesis of (2*R*)-methoxymalonyl-ACP. The FkbH-like Mad14 likely catalyses formation of glyceryl-ACP from 1,3 bisphosphoglycerate and the acyl-carrier protein Mad12. Glyceryl-ACP then undergoes two oxidation steps, catalysed by the dehydrogenases Mad11 and Mad13 to form (2*R*)-hydroxymalonyl-ACP, which is methylated by Mad15.

3.1.6 Tetrahydrofuran ring formation

The *mad* cluster encodes an epoxidase (MadC) and an epoxide hydrolase (MadB), homologues of which are present in the BGCs of many other tetrahydrofuran-containing polyethers including monensin[193], nanchangmycin[245], nigericin[246], lasalocid[194], and tetronomycin[111]. In a mechanism first proposed for monensin[247], the epoxidase and epoxide hydrolase catalyse a regio- and stereospecific oxidation cyclisation to form one or multiple rings. The energetically favourable product of such a reaction is a tetrahydrofuran ring, formed though an *exo* cyclisation, although tetrahydropyran rings may also be formed by the less favourable *endo* cyclisation route[194]. It is likely that MadC catalyses epoxidation of the C24-C25 and C28-C29 *E* double bonds. MadB then catalyses opening of the two epoxide rings to form two tetrahydrofuran rings (**Figure 3.13**). The timing of tetrahydrofuran formation in polyether is not certain, though evidence from other polyether pathways suggests it may occur whilst the intermediate is still bound to the PKS[196,246].

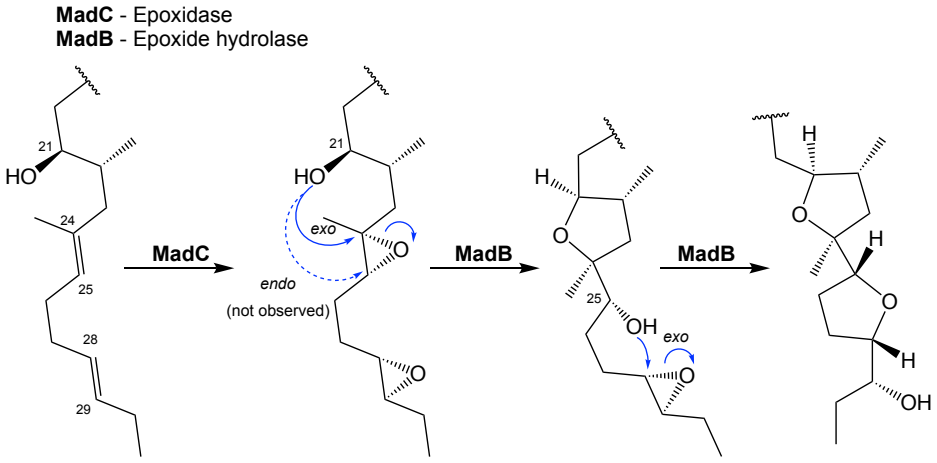


Figure 3.13 Proposed biosynthesis of the tetrahydrofuran rings in tetromadurin. MadC catalyses the regio- and stereospecific epoxidation of the C24-C25 and C28-C29 *E* double bonds. Next, the MadB epoxide hydrolase catalyses two sequential *exo* ring-closures to form the two tetrahydrofuran rings.

3.1.7 Hydroxylation of C36 and C38

Two cytochrome P450 enzymes are encoded in the *mad* cluster. One of these, Mad29, strongly resembles Tsn12 and Tmn14—the cytochrome P450s from the tetronasin and tetronomycin BGCs, respectively[111]. Deletion of *tsn12* in *S. longisporoflavus* led to the production of a tetronasin intermediate lacking the C30 hydroxyl group (Yuhui Sun, *unpublished*). Likewise, Tmn14 catalyses hydroxylation of the equivalent carbon (C28) in tetronomycin[111]. By analogy, Mad29 is therefore proposed to catalyse hydroxylation of C36 in tetromadurin (**Figure 3.14**), leaving the second cytochrome P450, Mad30, to catalyse hydroxylation of C38. No ferredoxin gene is present in the *mad* cluster, so presumably one encoded elsewhere in the genome is used to regenerate the cytochrome cofactor.

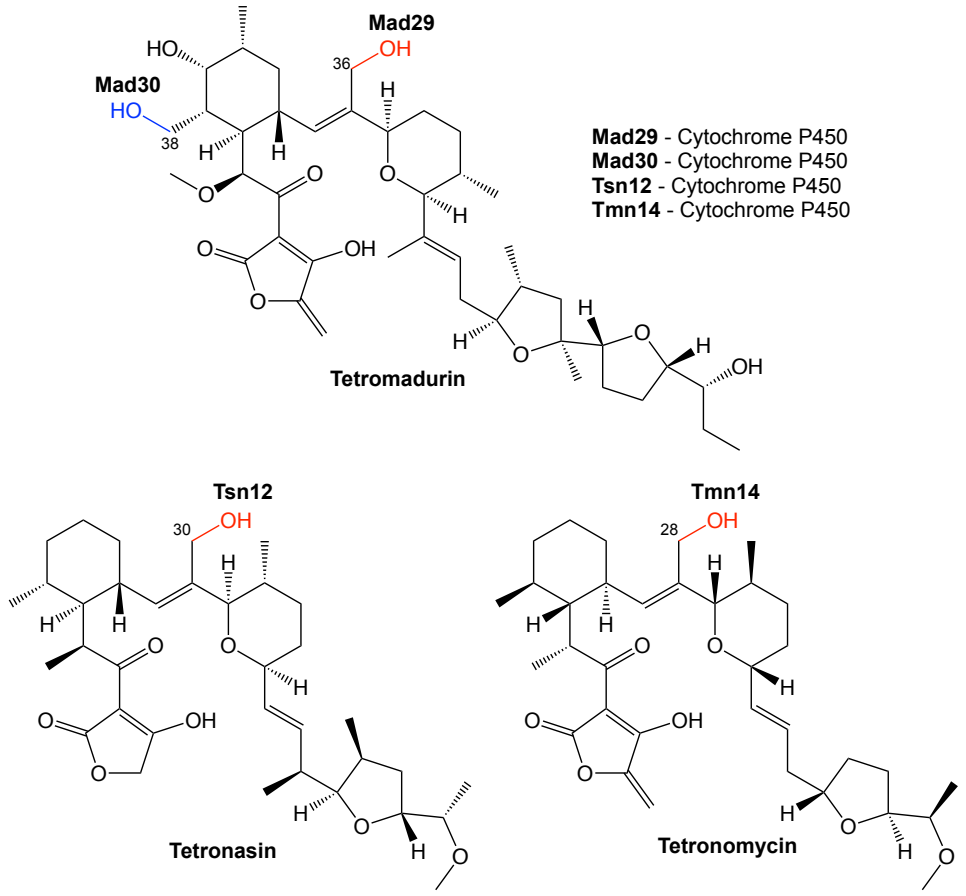


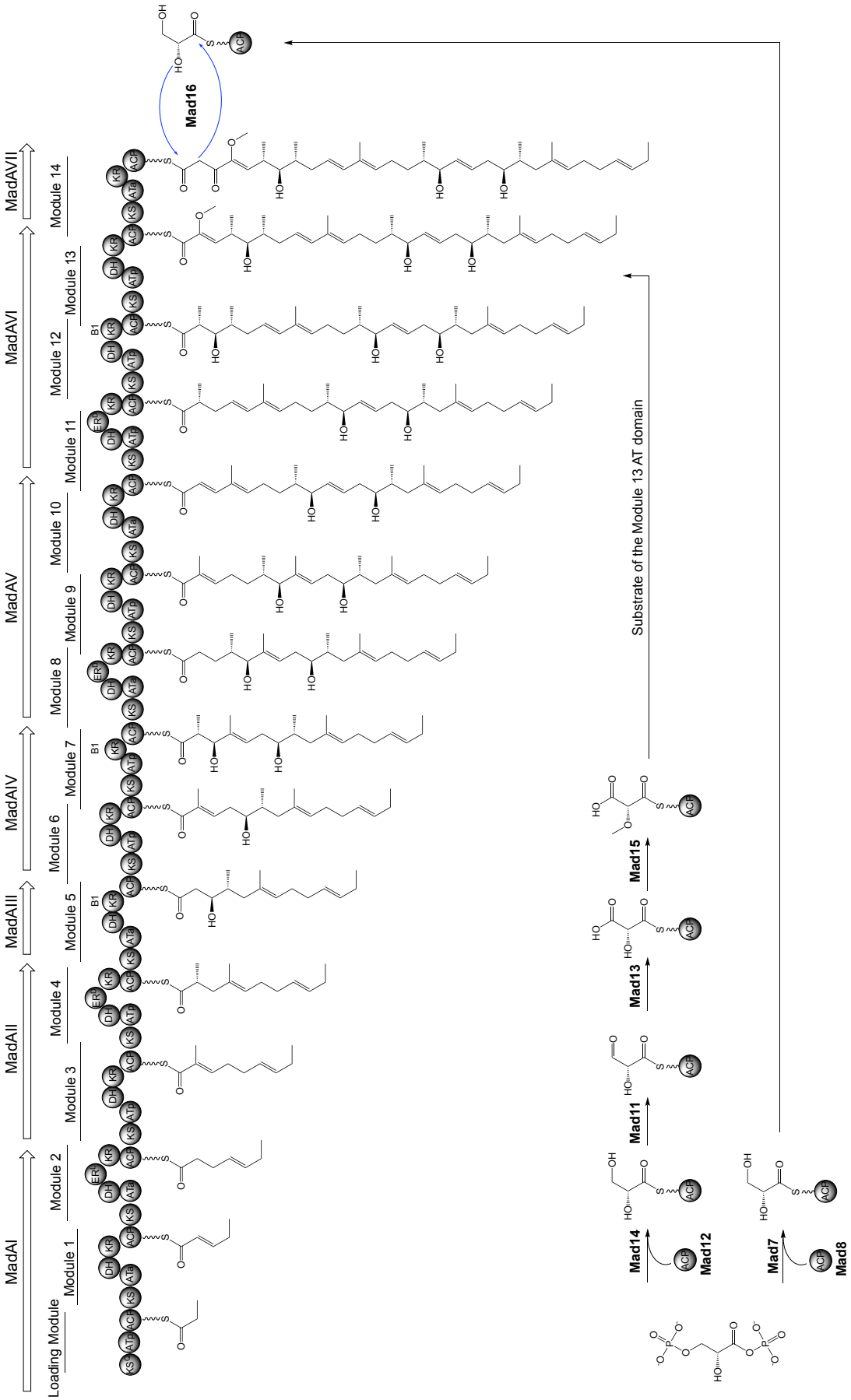
Figure 3.14 Proposed cytochrome P450-catalysed hydroxylations in tetromadurin biosynthesis. Mad29 is a homologue of the cytochrome P450 enzymes Tsn12 and Tmn14 present in the tetronasin and tetronomycin BGCs, respectively. It is therefore likely that Mad29 catalyses hydroxylation of C36, leaving Mad30 to hydroxylate C38.

3.1.8 Other genes in the *mad* cluster

The *mad* cluster contains other genes likely involved in transcriptional regulation of the *mad* cluster and export of tetromadurin itself. On one end of the *mad* cluster the genes *mad1* and *mad2* encode predicted SARP[248,249] and LuxR[250] transcriptional regulators, respectively. Also within the *mad* cluster is the gene *mad27* encoding a putative MarR (multiple antibiotic resistance regulator)-like transcriptional regulator. First described in *E. coli*, *marR* is a transcriptional regulator of several genes that confer antibiotic resistance[251]. Adjacent to *mad*27 are the genes *mad25* and *mad26* which encode an ABC transporter and an ABC transporter ATP-binding protein, respectively, both of which are likely involved in exporting tetromadurin from the cytoplasm[252]. Mad27 may induce expression of *mad25* and *mad26* in response to tetromadurin, resulting in self-resistance[198, 253].

A gene for a 4'-phosphopantetheinyl transferase (*mad3*) is also present, required for activating ACP domains by attaching a 4'-phosphopantetheine prosthetic group. Adjacent to *mad3* is a phosphoesterase *mad4*, homologues of which are found in other polyketide BGCs[111,112] and which may hydrolyse ACP-bound 4'-phosphopantetheine groups[112,254,255]. A stand-alone gene for a type II thioesterase is also present, likely having an "editing" role in hydrolysing PKS active sites containing mis-acylated intermediates[256,257]. Of the remaining genes in the *mad* cluster, *mad20* encodes a hypothetical protein with no characterised homologues. The putative functions of the final two genes, *mad10* and *mad31* were initially unclear prior to the analysis performed in the next section[157,161].

The bioinformatic analysis of the *mad* cluster strongly supported that it was responsible for tetromadurin biosynthesis. Experimental evidence confirming this conclusion will be presented in **Chapter 4**. Meanwhile, the proposed biosynthetic pathway to tetromadurin is shown in **Figure 3.15**.


MadAI
MadAII
MadAIII
MadAIV
MadAV
MadAVI
MadAVII
Loading Module
Module 1
Module 2
Module 3
Module 4
Module 5
Module 6
Module 7
Module 8
Module 9
Module 10
Module 11
Module 12
Module 13
Module 14
Mad16
Substrate of the Module 13 AT domain
Mad14
Mad12
Mad11
Mad13
Mad15
Mad7
Mad8

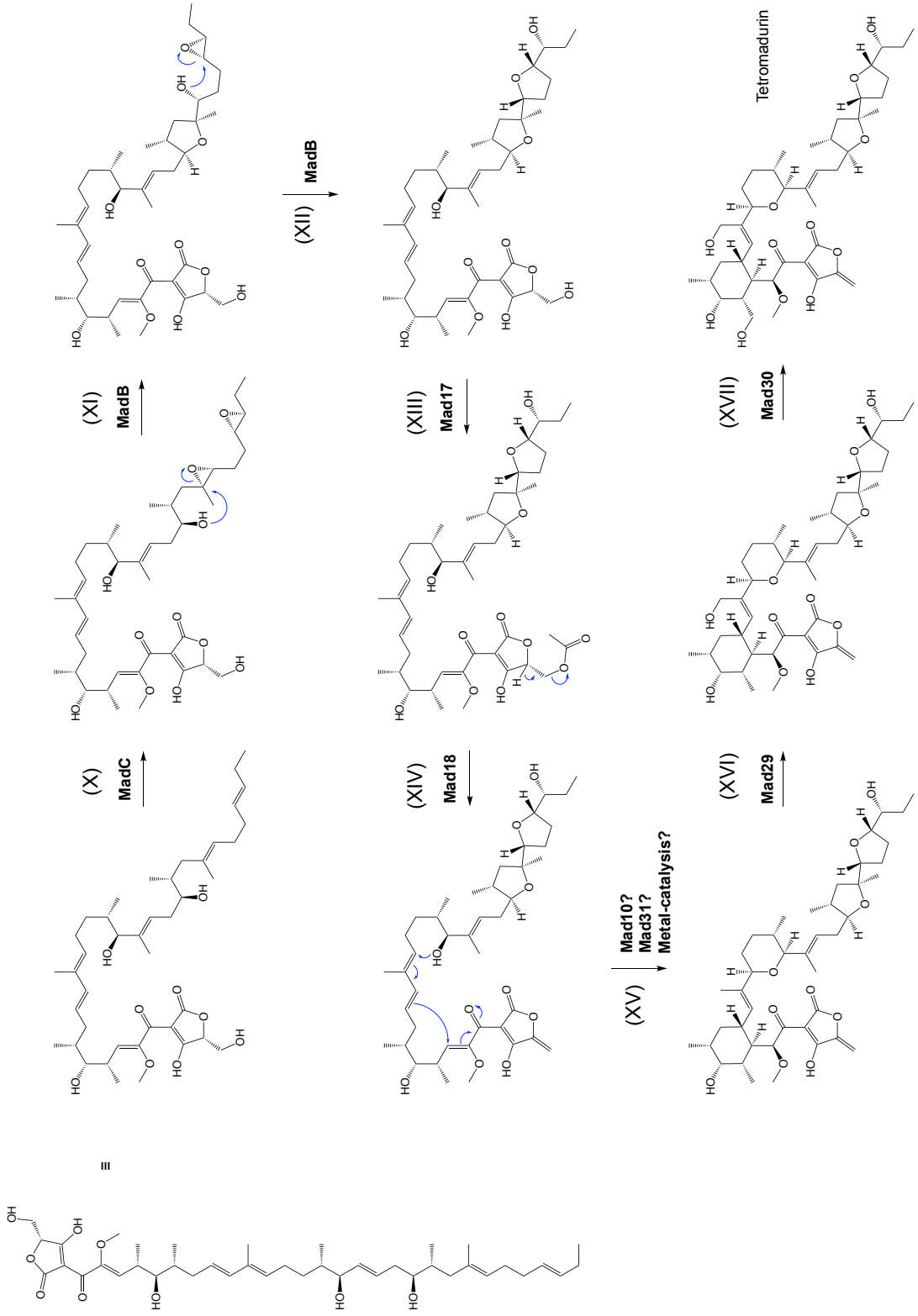

(XII)
MadB
(XI)
MadB
(X)
MadC
(XIII)
Mad17
(XIV)
Mad18
(XV)
Mad10?
Mad31?
Metal-catalysis?
(XVI)
Mad29
(XVII)
Mad30
Tetromadurin
III

Figure 3.15 Proposed biosynthetic pathway to tetromadurin. The order of post-PKS modifications could be different than presented here (I) The seven PKS enzymes (MadAI-MadAVIII) and predicted linear tetromadurin intermediates of the *mad* cluster. ATa domains contain the amino acid motifs associated with malonyl-CoA incorporation. ATp domains contain the amino acid motifs associated with (2*S*)-methylmalonyl-CoA incorporation. ER^D domains are predicted to create a D-configured α-methyl group. B1 KRs are predicted to create D-configured α-methyl and β-hydroxyl groups. The KS domain in MadAI is a KS^Q domain. (II) The key biosynthetic precursor 1,3-bisphosphoglycerate, involved in both tetronate formation and (2*R*)-methoxymalonyl-ACP biosynthesis. (III) Formation of glyceryl-ACP by the FkbH-like enzyme Mad14 and the ACP Mad12. (IV) Oxidation of glyceryl-ACP by the dehydrogenase Mad11 to form 2-hydroxy-3-oxopropionyl-ACP. (V) Oxidation of 2-hydroxy-3-oxopropionyl-ACP by the dehydrogenase Mad13 to form hydroxymalonyl-ACP. (VI) Methylation of hydroxymalonyl-ACP by the *O*-methyltransferase Mad15 to form (2*R*)-methoxymalonyl-ACP. (VII) The (2*R*)-methoxylmalonyl-ACP is incorporated by the MadAVI_AT_13. (VIII) Synthesis of a second pool of glyceryl-ACP by FkbH-like Mad7 and ACP Mad8. (IX) Glyceryl-ACP is a substrate of the FabH-like protein Mad16 which catalyses tetronate formation and concomitant chain release from the PKS. (X) In the first step of tetrahydrofuran formation, MadC catalyses epoxidation of the C24-C25 and C28-C29 *E* double bonds. (XI-XII) Epoxide hydrolase MadB catalyses formation of the two tetrahydrofuran rings through cascade epoxide ring-opening; (XIII) Acetylation of the C41 hydroxyl group catalysed by Mad17. (XIV) Elimination by Mad18 forms the C40-C41 exocyclic double bond. (XV) The formation mechanism of the cyclohexane and tetrahydropyran is unknown, but metal-catalysis is a possibility (as proposed for tetronasin[174,195]). (XVI) Hydroxylation of C36 catalysed by Mad29. (XVII) Hydroxylation of C38 by Mad30.

3.2 [4+2] cyclase homologues in the BGCs of tetronasin, tetronomycin, and tetromadurin

3.2.1 VstJ-homologues

Now that the sequence *tsn, tmn,* and *mad* clusters were all available, BLAST[229] could be used to search them for putative cyclase enzymes. Homologues of *tmn8*/*vstJ* were identified in both the *tsn* cluster (*tsn15*) and the *mad* cluster (*mad31*) (**Figure 3.16**) (**Appendix Note 1**). However, alignment of the protein sequences of the polyether tetronate and spirotetronate/tetramate VstJ-homologues indicated that there is only modest sequence identity (typically between 20-28%) between any the enzymes (**Figure 3.17**).

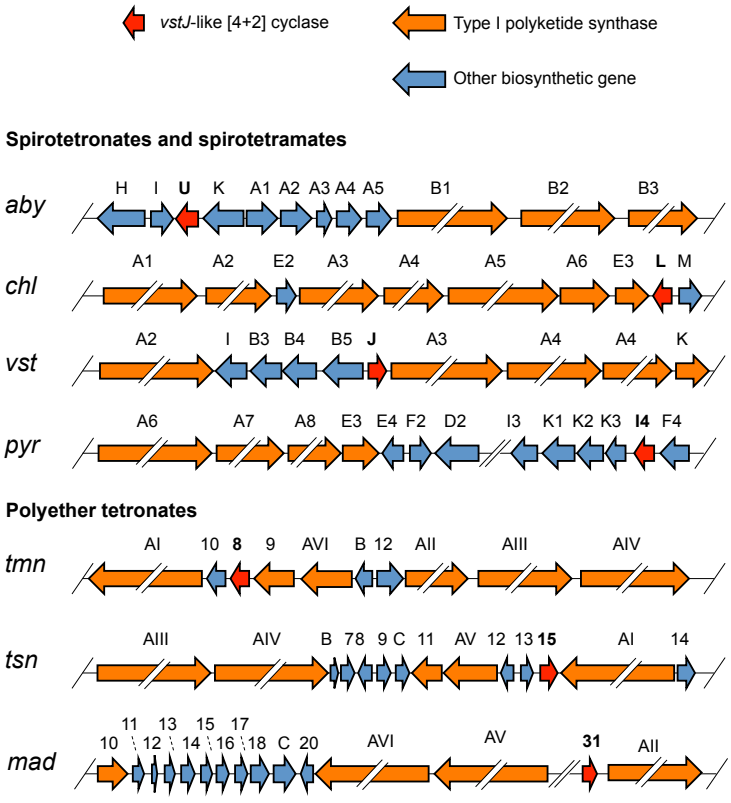


Figure 3.16 *vstJ-type* [4+2] cyclases in the biosynthetic gene clusters of tetronasin and tetromadurin. The BGCs of the spirotetronates versipelostatin (*vst*), abyssomicin (*aby*), chlorothricin (*chl*), the spirotetramate pyrroindomycin (*pyr*), and the polyether tetronates tetronomycin (*tmn*), tetronasin (*tsn*), and tetromadurin (*mad*) all encode a VstJ-like [4+2] cyclase (highlighted in red). Only partial BGCs are shown and genes are not drawn to scale.

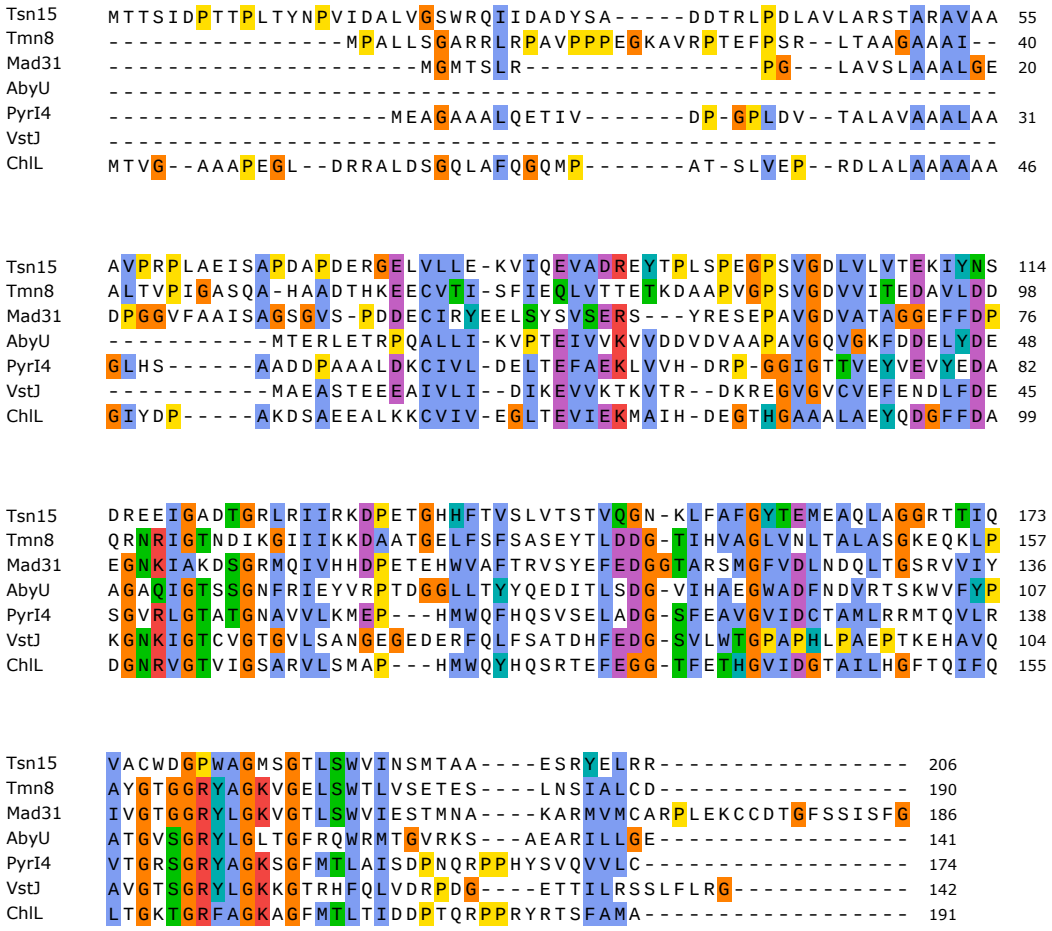


Figure 3.17 Alignment of VstJ-type [4+2] cyclases. Protein sequence alignment of [4+2] cyclases encoded in the biosynthetic gene clusters of the polyether tetronates tetronasin (Tsn15), tetronomycin[111] (Tmn8), and tetromadurin (Mad31); the spirotetronates abyssomicin (AbyU), versipelostatin (VstJ), chlorothricin (ChlL) and the spirotetramate pyrroindomycin (PyrI4). Similar residues in an alignment are coloured accordingly: Blue, hydrophobic; Green, polar; Purple, negative; Red, positive; Yellow, proline; Orange, glycine. Alignment was made using ClustalOmega[227].

Additional support that Tmn8, Tsn15, and Mad31 are genuine homologues of the VstJ-like [4+2] cyclases, despite the low sequence similarity, was obtained using the Phyre2 server. Phyre2 is an online tool that predicts protein structure by matching its primary sequence and predicted secondary structure to previously solved protein structures[258]. Phyre2 analysis of Tsn15, Tmn8, Mad34, and VstJ predicted that all shared an allene oxide cyclase barrel fold (**Figure 3.18**). The solved protein structures of PyrI4[159] and AbyU[158] have since confirmed the relationship to allene oxide cyclases.

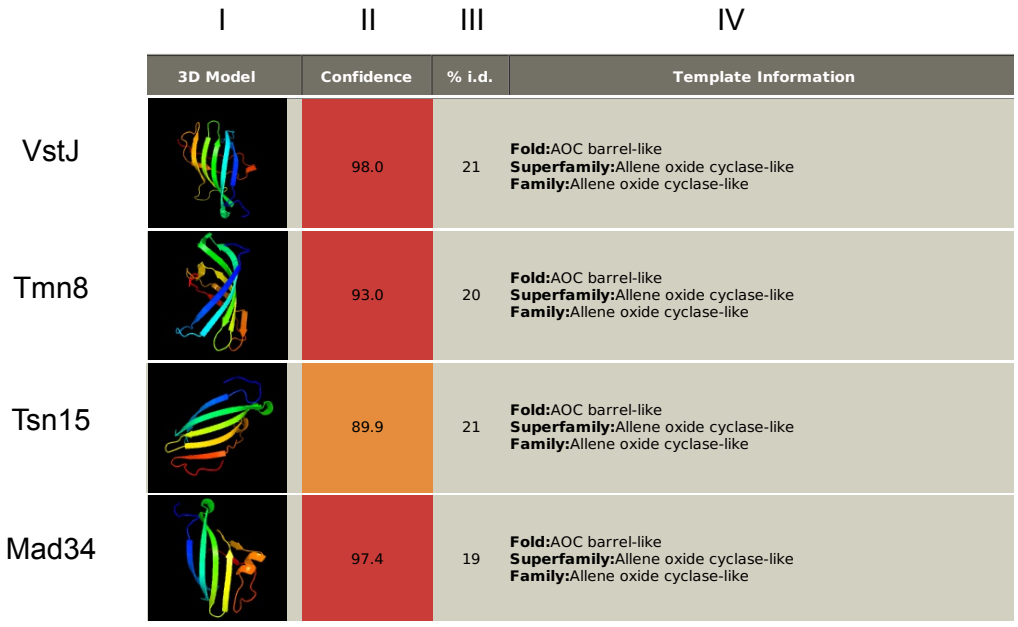


Figure 3.18 Phyre2 structural prediction for VstJ, Tmn8, Tsn15, and Mad34. The primary sequences of VstJ, Tmn8, Tsn15, and Mad34 were analysed using the Phyre2 software[258]. (I) Predicted 3D model of the protein, (II) % confidence in the 3D model, (III) % identity between the query sequence and the allene oxide cyclase fold library (IV) Information on the 3D template used.

Taken together, BLAST and Phyre2 results confirmed that VstJ-like [4+2] cyclase homologues are encoded in all three polyether tetronate BGCs. The function of these enzymes was unknown, but their conservation suggested an essential biosynthetic role.

3.2.2 PyrE3 homologues

In addition to VstJ-like [4+2] cyclases, a second class of [4+2] cyclases was also identified for the first time in polyether tetronate BGCs. This second class are homologues of PyrE3, the [4+2] cyclase that catalyses dialkyldecalin formation in pyrroindomycin biosynthesis[161] (**Figure 1.33**). Using PyrE3 as a BLAST query identified homologues encoded in all three polyether tetronate BGCs: Tmn9 from the *tmn* cluster, Tsn11 from the *tsn* cluster, and Mad10 from the *mad* cluster (**Figure 3.19**) (**Appendix Note 2 and 3**). In contrast to the VstJ homologues, the PyrE3 homologues shared a higher degree of sequence identity, ca. 40-45%. Like PyrE3, BLAST analysis of Tmn9, Tsn11, and Mad10 initially annotated them as FAD-dependent monooxygenases of unknown biosynthetic function[111].

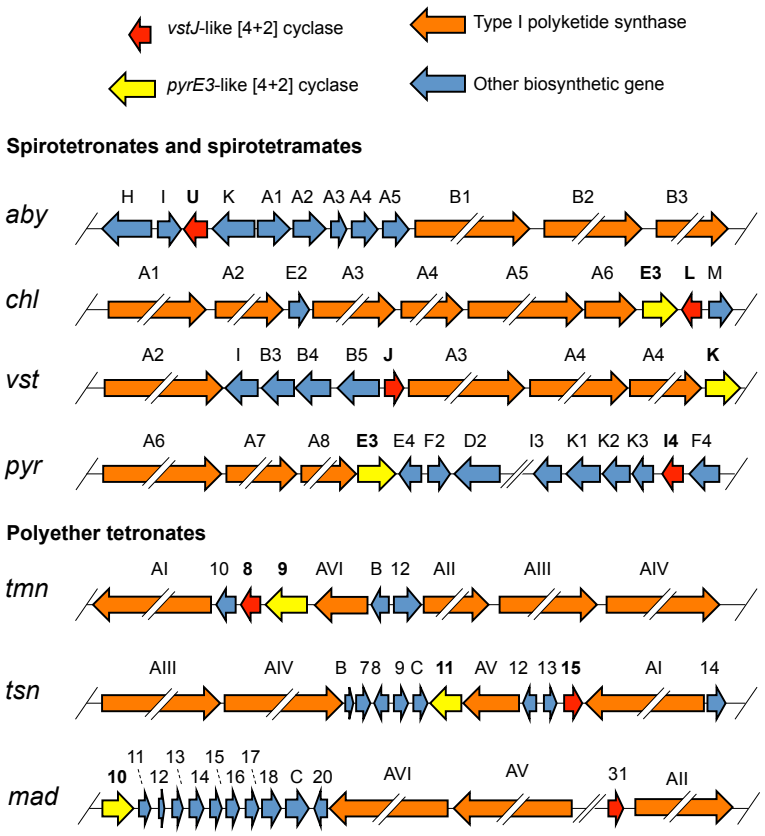


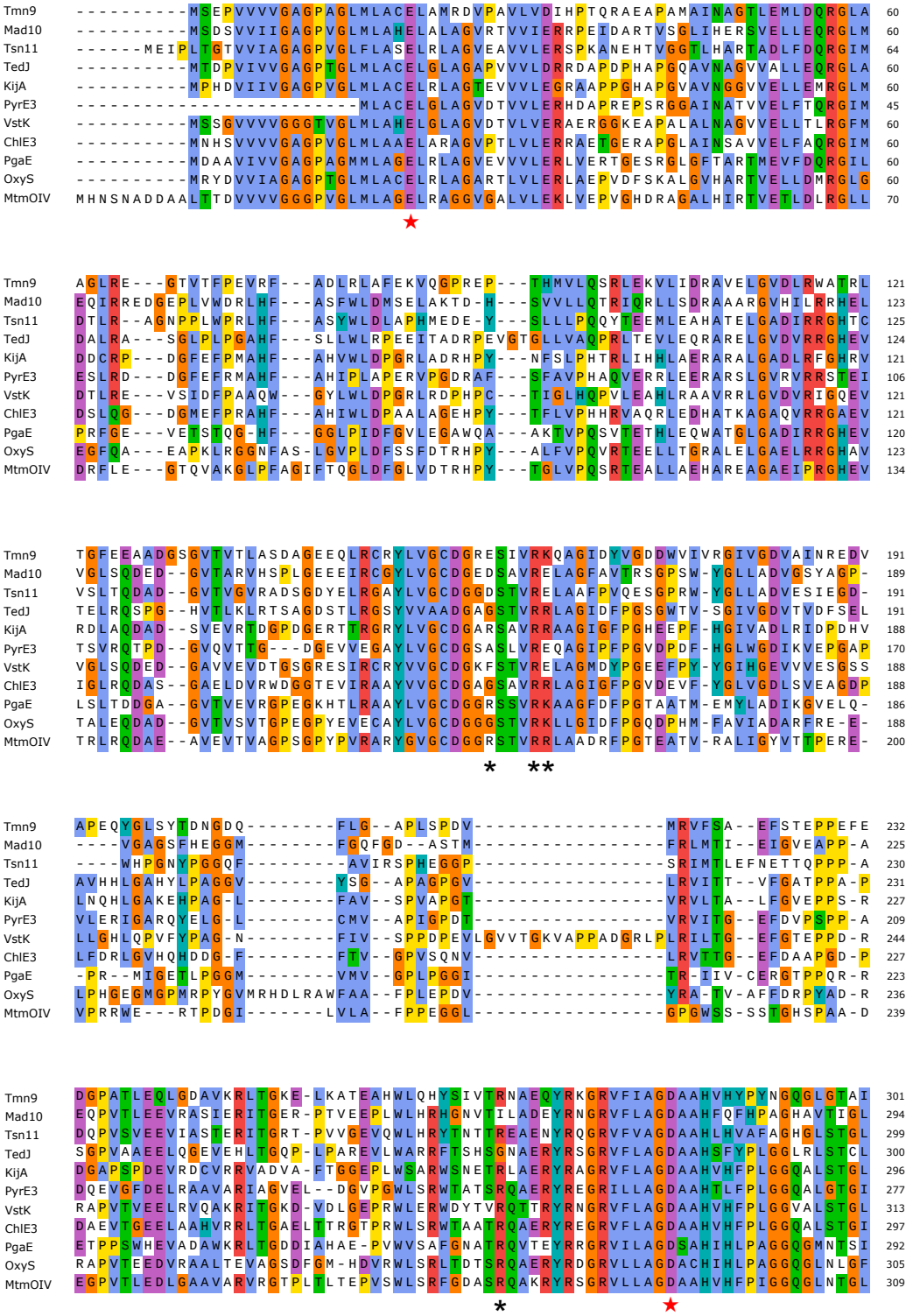
Figure 3.19 *pyrE3*-like [4+2] cyclases in the biosynthetic gene clusters of polyether tetronates. The BGCs of the spirotetronates versipelostatin (*vst*), chlorothricin (*chl*), the spirotetramate pyrroindomycin (*pyr*), and the polyether tetronates tetronomycin, tetronasin, and tetromadurin all encode a PyrE3-like [4+2] cyclase (highlighted in yellow) and a VstJ-like [4+2] cyclase. Only partial BGC are shown and genes are not to scale.

Tmn9 had even been previously deleted in *S. sp* NRRL 11266, demonstrating that it was essential for tetronomycin biosynthesis[111]. Further, inspection of a maximum-likelihood phylogenetic tree revealed that Tmn9, Mad10, and Tsn11 form a clade together but are more related to the PyrE3-like [4+2] cyclases than to their *bona fide* FAD-dependent monooxygenases homologues (**Figure 3.20**).

Figure 3.20 Maximum likelihood phylogenetic tree of PyrE3 homologues. A maximum likelihood phylogenetic tree was created for the diverse family of PyrE3 homologues, including: [4+2] cyclases from spirotetronate/tetramate pathways; FAD-dependent monooxygenases; and homologues from polyether tetronate pathways. 1000 bootstrap replicates were performed.

Despite this, Tmn9, Mad10, and Tsn11 all likely bind to FAD as they contain the negatively charged residues that are required in their OxyS/MtmOIV-like monooxygenase homologues required for FAD binding[163,259] (**Figure 3.21**). Although PyrE3 also binds an FAD cofactor, it was shown to be unable to oxidise NADPH, indicating that the FAD does not play a redox role (as it would in a true FAD-dependent monooxygenase)[163,259,260]. Work on MtmOIV identified four positively charged residues crucial for the proper binding of NADPH (R169, R173, R174, R277). Although mutating these residues to alanine did not significantly affect the NADPH K_m binding constant, it did decrease the NADPH k_{cat} value, likely by distorting NADPH binding and preventing effective FAD reduction[260]. PyrE3 contains mutations at several of these positions, suggesting its inability to oxidise NADPH

is likely due to distorted NADPH binding[163]. Mutations in these residues are also found in other PyrE3 homologues with [4+2] cyclase activity such as ChlE3, TedJ, and VstK, consistent with these enzymes being [4+2] cyclases rather than monooxygenases[163]. Tmn9, Mad10, Tsn11 are also highly variable in these residues, suggesting they also cannot effectively bind to NADPH (Figure 3.21). Exceptions exist however, as the PyrE3 homologue from the spirotetronate kijanimicin, KijA, contains arginine at each of the four loci, indicating it may retain the ability to oxidise NADPH. Nevertheless, that the majority of known PyrE3-type [4+2] cyclases lack arginine at these sites provides evidence against Tmn9, Mad10, and Tsn11 having FAD-dependent monooxygenase activity and points in the direction of them being cyclases.


Tmn9
Mad10
Tsn11
TedJ
KijA
PyrE3
VstK
ChlE3
PgaE
OxyS
MtmOIV

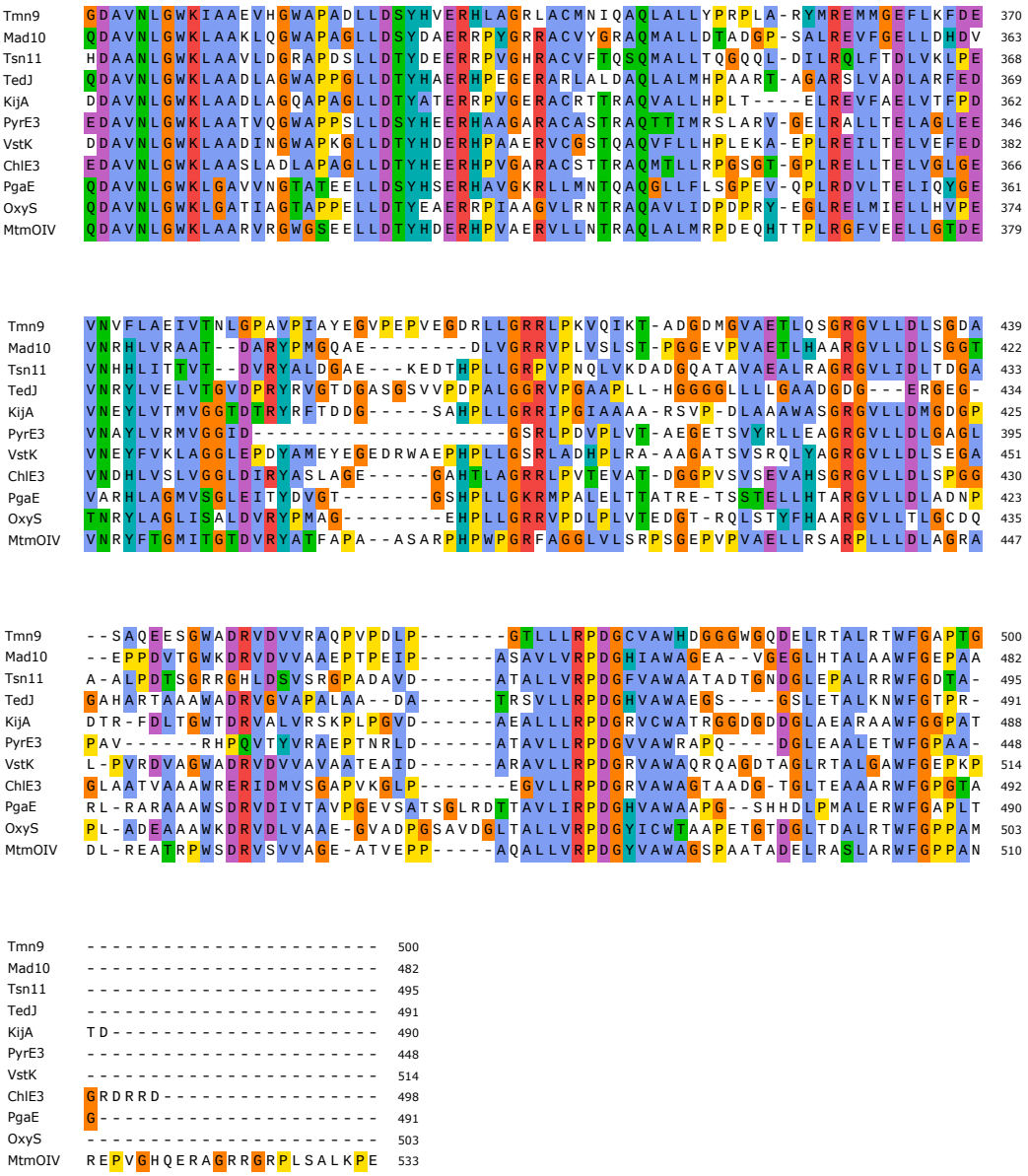


Figure 3.21 Alignment of PyrE3-type [4+2] cyclases. Protein sequence alignment of the PyrE3-type [4+2] cyclases encoded in the biosynthetic gene clusters of the polyether tetronates tetronomycin (Tmn9), tetronomadurin (Mad10), and tetronasin (Tsn11); The [4+2] cyclase TedJ from the tetrodecamycin BGC[118], The [4+2] cyclase KijA from the kijanimicin BGC[115]; The [4+2] cyclase PyrE3 from the pyrroindomycin BGC[161]; The [4+2] cyclase VstK from the versipelostatin BGC[157], The [4+2] cyclase ChlE3 from the chlorothricin BGC[116,161]; The monooxygenase PgaE from gaudimycin C biosynthesis[261]; The monooxygenase OxyS from oxytetracycline biosynthesis[259]; the monooxygenase MtmOIV from mithramycin biosynthesis[260]. Red stars (★) indicate amino acids involved in binding FAD[259]. Black asterisks (*) indicate the locations of the four arginine residues in MtmOIV involved in NADPH binding[260]. Alignment made using ClustalOmega[227]

3.3 Chapter 3 Discussion

The initial automated analysis of the assembled genome of *A. verrucosispora* revealed the *mad* cluster as the obvious candidate for tetromadurin biosynthesis. The genes present in the *mad* cluster were a near-perfect fit for both synthesising the tetromadurin carbon backbone and performing the necessary post-PKS tailoring modifications. The single unexplained feature of the PKS is the origin of the C7 hydroxyl group, which is formed despite the KR of module 12 lacking its essential active site tyrosine and several NADPH binding site residues. It may be that the KS, AT, and ACP domains of module 12 catalyse an extension cycle using (2*S*)-methylmalonyl-CoA, which is reduced only by the KR domain of an adjacent module. Domains acting externally to the PKS module they are located in have been proposed for other polyketide biosynthesis pathways[64–67]. Likewise, some domains within iteratively acting modules only function during one of the iterative cycles, as was demonstrated for the ER domain of azalomycin F[68]. The KR domain of either module 11 or 13 could therefore be responsible for forming the C7 hydroxyl. Like the ER domain in azalomycin F biosynthesis, the other reductive domains in these modules may be switchable between active and inactive states, preventing reduction beyond the hydroxyl[266]. The final possibility is that, despite appearing inactive, the KS of module 12 really is a functional domain. To investigate this in the future, the domain could be independently purified to test its activity *in vitro,* as has been done for other KR domains[267].

Having the sequence of the *mad* cluster provided several useful insights into polyether tetronate biosynthesis. First, both the *tmn* and *tsn* clusters contain additional flanking PKS genes (tmn/tsnD genes) not necessary for the biosynthesis of tetronomycin and tetronasin, respectively[111]. The *mad* cluster, however, does not have these flanking PKS genes (nor are they found anywhere on the genome), confirming that they are not a required feature of polyether tetronate BGCs. The glycerate utilisation operon in the *mad* cluster also further highlights the uniqueness of the hydroxypyruvate dehydrogenase genes in the *tsn* cluster (Yuhui Sun, *unpublished results*).

The second key finding of this chapter was that the gene clusters of all three polyether tetronates each contain two different genes for [4+2] cyclase homologues. The conservation of these genes across all three clusters, paired with the functional evidence that Tmn9 is essential for tetronomycin production[111], strongly suggested that these

enzymes are essential for polyether tetronate biosynthesis. Tmn8, Tsn15, and Mad31 are homologues of the VstJ-like [4+2] cyclases that catalyse biosynthesis of the spirotetronate/spirotetramate moiety[157,158,161]. Meanwhile Tmn9, Mad10, and Tsn11 meanwhile are all homologues of the PyrE3-like [4+2] cyclases that catalyse dialkyldecalin formation in spirotetronate/spirotetramate natural products[118,161]. Considering that polyether tetronate biosynthesis is not predicted to require even one [4+2] cyclisation, let alone two, the role of these [4+2] cyclase homologues was a mystery. Phylogenetic analysis of Tmn9, Mad10, and Tsn11 indicated they are more related to [4+2] cyclases than to monooxygenases and lack the residues associated with NADPH binding[260], suggesting that they have a non-redox function.

In summary, the BGC of tetronasin, tetromadurin, and tetronomycin were each revealed to contain two overlooked cyclase homologues. A reasonable hypothesis at this stage was that these cyclase homologues might be involved in cyclohexane and tetrahydropyran formation. Experiments to test this hypothesis directly are described in the next chapter.

Chapter 4: *in vivo* studies on polyether tetronate [4+2] cyclase homologues

Formal proof that the polyether tetronate homologues of VstJ and PyrE3 are essential for biosynthesis required specific deletion of each gene and the demonstration that production was lost. Another strong motivation to attempt these deletion experiments was to assess if whether by blocking production of the tetronate antibiotic any intermediates or shunt products might instead be produced, which might shed light on the role of the deleted gene. In previous work on the *tmn8* mutant[111] no such metabolite was reported, but this is a low-yielding strain and compounds may therefore have been overlooked. Further encouragement for such experiments was provided by the precedent that deletion of *vstJ* and *pyrE3* homologues resulted in the mutant strain producing a late-stage precyclisation intermediate[157,161]. The timing of ring formation may also influence the likelihood of an intermediate being detected. For instance, if ring formation occurs before chain release, as predicted in other polyethers[196,197], disrupting it may prevent the release of any polyketide intermediates from the PKS. To increase the chances of detecting metabolites in blocked mutants, it was decided to undertake deletions in all three polyether tetronate-producing strains.

4.1 Production of tetronomycin, tetromadurin, and tetronasin

Before undertaking gene deletions, it was confirmed that the available stocks of *S. longisporoflavus*, *S.* sp. NRRL 11266 and *A. verrucosispora* in the Leadlay lab were competent producers of their respective polyether tetronates.

4.1.1 Production of tetronomycin from *S.* sp. NRRL 11266

S. sp. NRRL 11266 was known to produce tetronomycin on a soy flour/D-mannitol medium (SFM)[111]. In the present study, *S.* sp. NRRL 11266 grew on SFM medium as dry, pale colonies that darkened into a brown colour before forming white spores. Production of tetronomycin was also detected from *S.* sp. NRRL 11266 in these conditions (**Figure 4.1**).

4.1.2 Production of tetromadurin from *A. verrucosispora*

For the tetromadurin producer *Actinomadura verrucosispora*, a range of different mediums was tested to find one suitable for tetromadurin production. *A. verrucosispora* was found to sporulate and produce tetromadurin on both oatmeal and ISP4 media (**Figure 4.2**). The mass and UV chromophore of tetromadurin were the same as previously reported[168].

4.1.3 Production of tetronasin from *S. longisporoflavus*

Streptomyces longisporoflavus 83E6 is a tetronasin overproducer developed by ICI Pharmaceuticals in the 1980s, shown in this laboratory to produce well on a dextrin-containing medium[208,209]. Here, *Streptomyces longisporoflavus* 83E6 grew on TWM medium as a pale, non-sporulating mycelia. Sporulation of *S. longisporoflavus* 83E6 was not observed on any growth medium tested. Tetronasin production was also readily detected when *S. longisporoflavus* was grown in tryptic soy broth supplemented with dextrin (tsn-medium-B) (**Figure 4.3**). Confirmation of its identity was obtained by comparison of its MS^3 fragmentation pattern with an authentic tetronasin sample[268] (**Appendix Figure 1**).

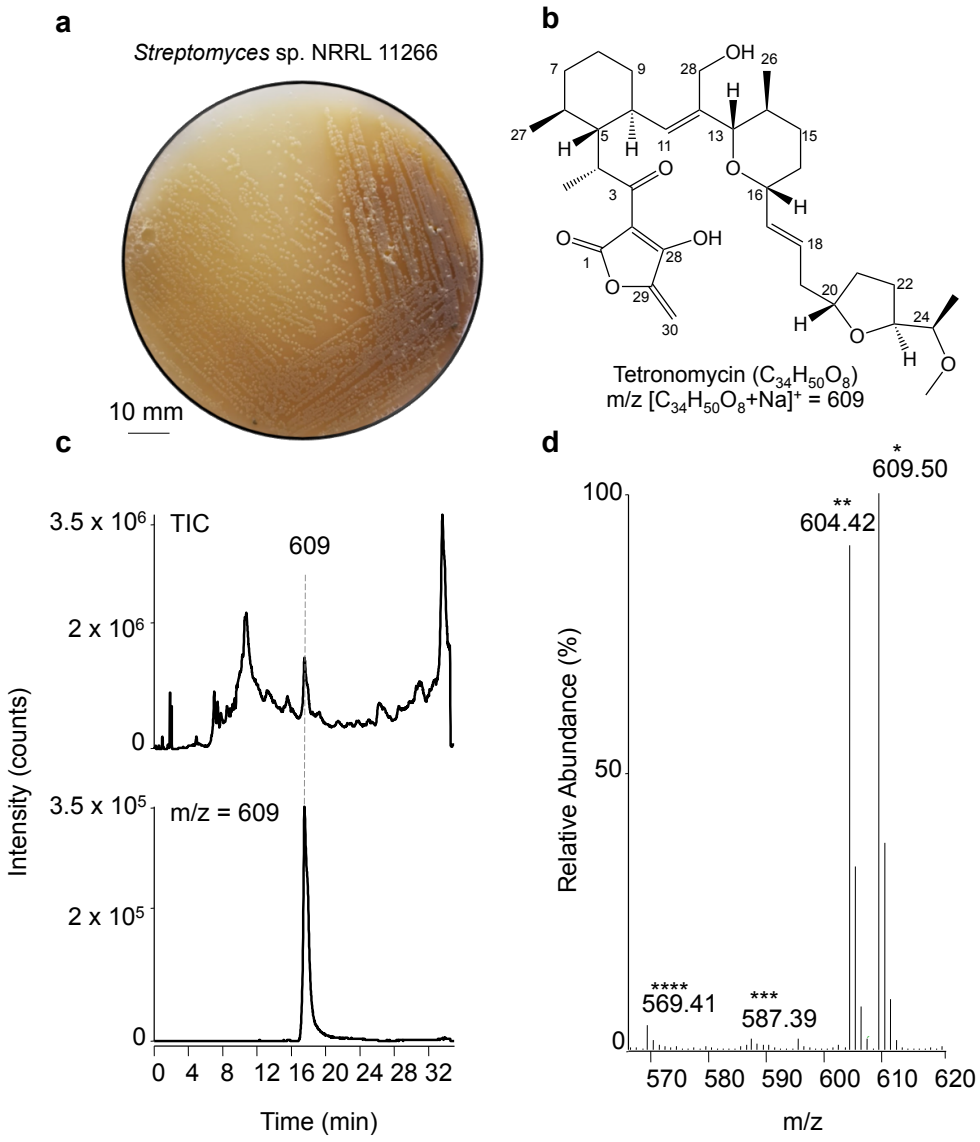


Figure 4.1 Production of tetronomycin by *Streptomyces* sp. NRRL 11266. a, Growth of *S.* sp. NRRL 11266 on SFM. **b,** Structure of tetronomycin ($C_{34}H_{50}O_8$) m/z $[C_{34}H_{50}O_8+Na]^+$ = 609.80. **c,** Top: total ion current (TIC) chromatogram of *Streptomyces* sp. NRRL 11266 organic extract. A clear peak corresponding to tetronomycin was present at 17.5 min. Bottom: Extracted m/z = 609 chromatogram of *S.* sp. NRRL 11266. **d,** Mass spectrum of tetronomycin *$[C_{34}H_{50}O_8+Na]^+$; **$[C_{34}H_{50}O_8+NH_4]^+$; ***$[C_{34}H_{50}O_8+H]^+$; ****$[C_{34}H_{50}O_8+H-H_2O]^+$.

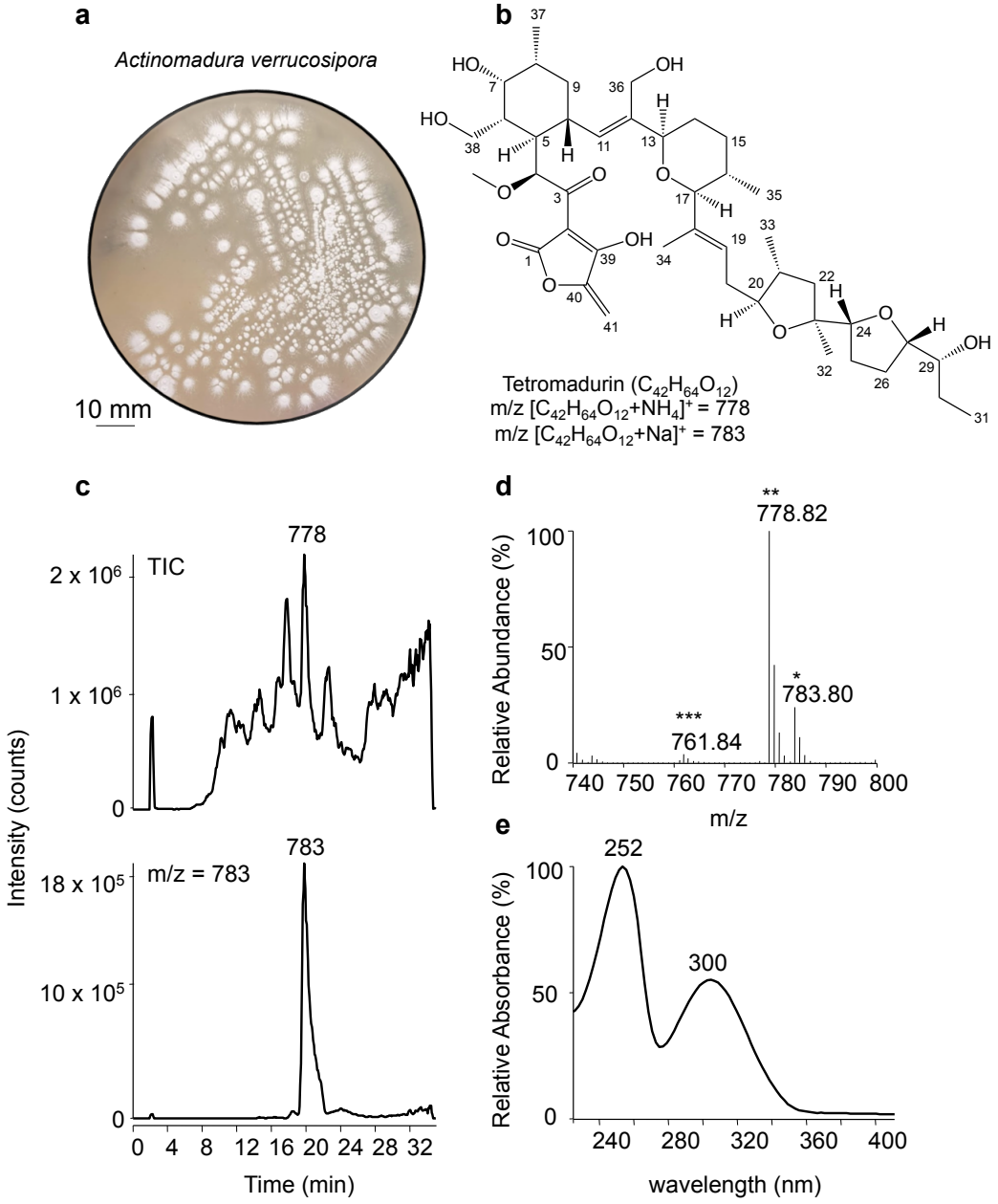


Figure 4.2 Production of tetromadurin by *Actinomadura verrucosispora*. a, *A. verrucosispora* grew on oatmeal agar as pale colonies that formed white hydrophobic spores after several days. **b,** Structure of tetromadurin ($C_{42}H_{64}O_{12}$), m/z $[C_{42}H_{64}O_{12}+NH_4]^+$ = 778; m/z $[C_{42}H_{64}O_{12}+Na]^+$ = 783. **c,** Top: total ion current (TIC) chromatogram of *A. verrucosispora* organic extract. A peak corresponding to tetromadurin $[C_{42}H_{64}O_{12}+NH_4]^+$ (m/z = 778.82) was visible on the TIC at 19.8 min. Bottom: extracted m/z = 783 chromatogram of the tetromadurin sodium ion adduct $[C_{42}H_{64}O_{12}+Na]^+$ **c,** Mass spectrum of the tetromadurin peak from the TIC spectrum. *$[C_{42}H_{64}O_{12}+Na]^+$; **$[C_{42}H_{64}O_{12}+NH_4]^+$; ***$[C_{42}H_{64}O_{12}+H]^+$. **e,** UV absorption spectrum of the tetromadurin peak. λ_{MAX} = 252 nm, 300 nm (MeOH), as previously reported[168].

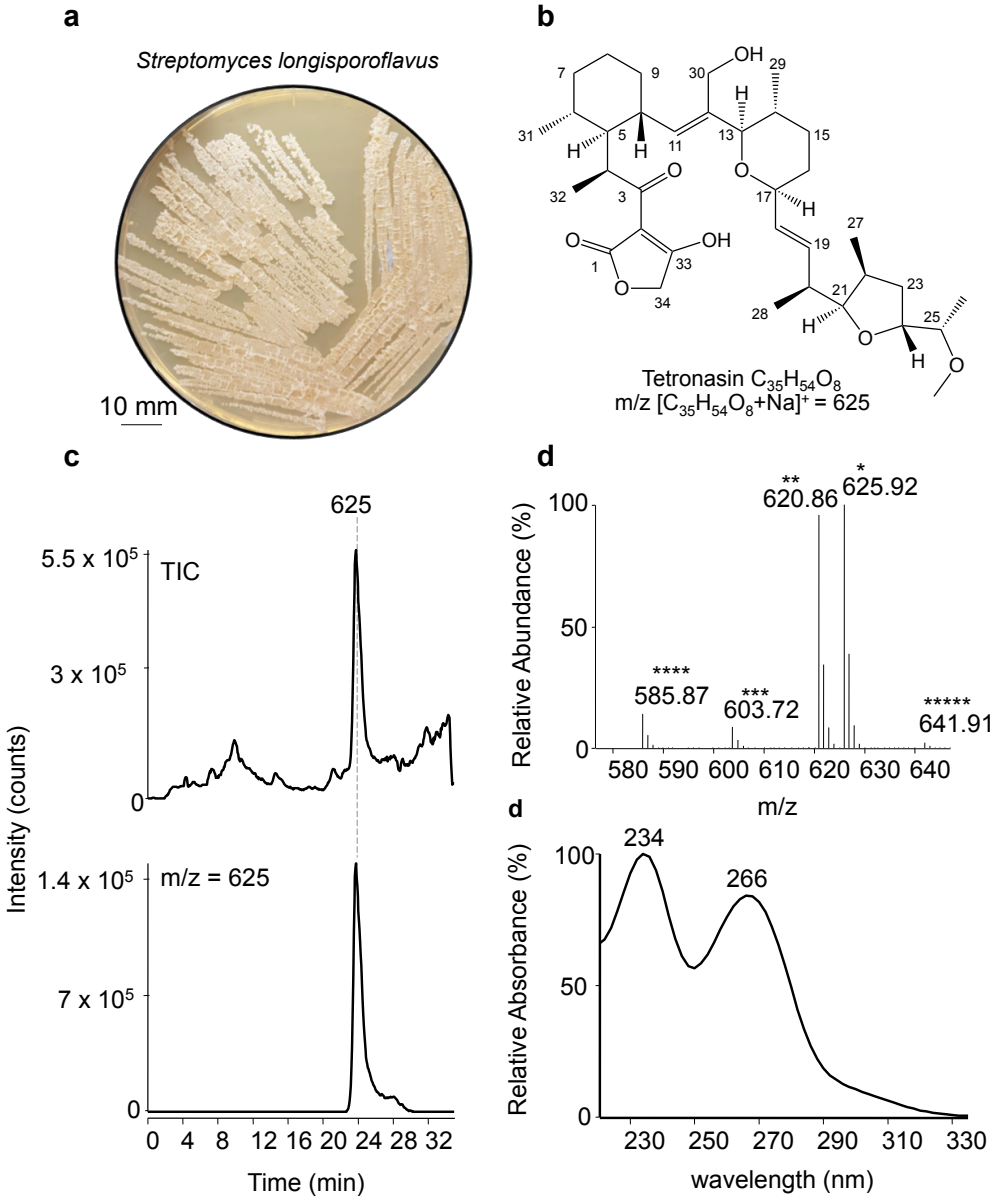


Figure 4.3 Production of tetronasin by *Streptomyces longisporoflavus*. a, *Streptomyces longisporoflavus* 83E6 grew on TWN medium as a pale, non-sporulating, mycelia. **b,** Structure of tetromadurin ($C_{35}H_{54}O_8$), m/z $[C_{35}H_{54}O_8+NH_4]^+$ = 620; m/z $[C_{35}H_{54}O_8Na]^+$ = 625. **c,** Top: total ion current (TIC) chromatogram of *S. longisporoflavus* organic extract. A peak corresponding to tetronasin $[C_{35}H_{54}O_8+Na]^+$ (m/z = 625) was visible on the TIC at 24.0 min. Bottom: Extracted chromatogram of the tetronasin m/z = 625 peak from the TIC. **c,** Mass spectrum of the tetronasin peak from the TIC. *$[C_{35}H_{54}O_8+Na]^+$; **$[C_{35}H_{54}O_8+NH_4]^+$; ***$[C_{35}H_{54}O_8+H]^+$; ****$[C_{35}H_{54}O_8+H-H_2O]^+$; *****$[C_{35}H_{54}O_8+K]^+$. **d,** UV absorption spectrum of the tetronasin peak: λ_{MAX} = 234 nm, 266 nm.

These preliminary fermentation experiments confirmed that all three polyether tetronates could be produced in detectable quantities from their respective producer strains. Any changes resulting from deletion of a *vstJ*- or *pyrE3-like* [4+2] cyclase should therefore be detectable by HPLC-MS.

4.2 Creation of polyether tetronate [4+2] cyclase deletion mutants

All gene deletions were performed by a homologous recombination-based protocol using the pYH7 plasmid. pYH7 was previously used to successfully delete genes in both *S. longisporoflavus* (Sun *et al*., *unpublished results*) and *S.* sp. NRRL 11266[111]. In each case, all or a significant portion of the target gene was deleted. To decrease the chances of the gene deletion having polar effects on the expression of adjacent genes, deletions were made in the coding frame so no frame shift was introduced[210]. As the *S.* sp NRRL 11266 Δtmn9[111] mutant had previously been analysed, it was not created again in this study. Regions of 2 kb directly flanking the target gene both upstream and downstream were amplified by PCR and assembled into pYH7 (an example of this, for the plasmid used to delete *tsn11*, is shown in **Figure 4.4**). The complete workflow for creating in-frame deletions in actinomycetes is presented in **Figure 4.6**.

A pYH7 deletion construct for each gene was delivered by intergeneric conjugation from *E. coli* into the appropriate actinomycete strain, where it integrated into the genome *via* a single crossover. The pYH7 plasmid contains the *acc(3)IV* gene, conferring resistance to apramycin. The pYH7-containing actinomycete "exconjugants" were grown for several generations in the absence of apramycin to cure the pYH7 plasmid by inducing a second crossing over event, which either reverted the strain to wild type or introduced the desired coding-frame deletion. Exconjugants that had undergone this second crossover were sensitive to apramycin again so could be identified by replica plating (**Figure 4.5**).

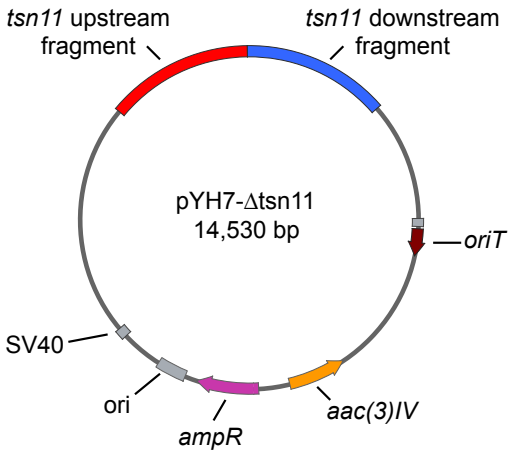


Figure 4.4 Plasmid map of the pYH7 construct used for the in-frame deletion of *tsn11* in *S. longisporoflavus*. The 2 kb upstream (red) and 2 kb downstream (blue) fragments were the sites of homologous recombination with the *S. longisporoflavus* chromosome, resulting in the coding-frame deletion of *tsn11*. *acc(3)IV*: apramycin resistance gene. *ampR*: ampicillin resistance gene (β-lactamase). *oriT*: origin of transfer, *ori*: origin of replication. SV40: SV40 polyadenylation signal.

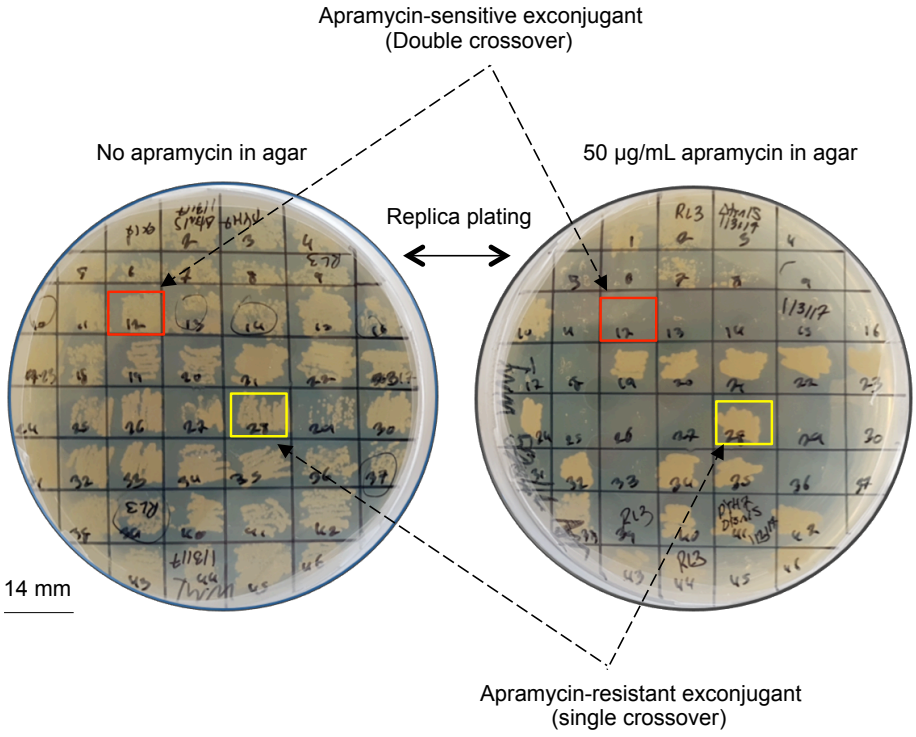


Figure 4.5 Replica plating of *S. longisporoflavus* exconjugants containing pYH7-Δtsn15. The apramycin-sensitive exconjugants had either undergone a second crossing over event with the pYH7-Δtsn15 plasmid, in which case *tsn15* had been deleted from the chromosome, or had reverted to wild type. The red boxes highlight an exconjugant that had lost the pYH7 plasmid and therefore regained apramycin sensitivity. The yellow boxes highlight an exconjugant that still contains the pYH7-Δtsn15 plasmid and is therefore apramycin resistance.

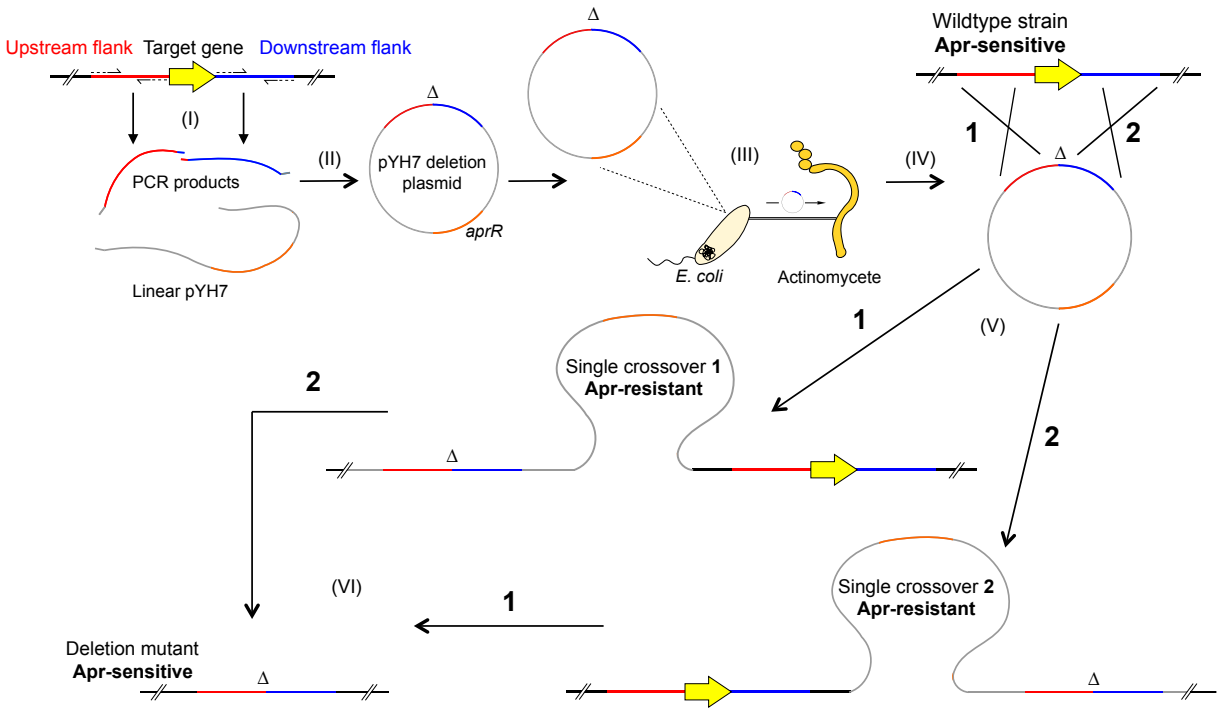


Figure 4.6 Workflow for the deletion of [4+2] cyclase homologues in polyether-tetronate producers. (I) 2 kb regions upstream and downstream of the target gene were amplified using PCR. (II) The two flanking regions were joined and ligated into the pYH7 plasmid using Gibson assembly. (III) The assembled pYH7 deletion plasmid was conjugated into the tetronate-producing actinomycete using an *E. coli* donor. (IV) Once conjugated, the flanking regions in the pYH7 plasmid could crossover with the homologous regions in the chromosome. (V) Either of the two flanking regions could crossover first, creating two alternative single crossover genotypes (1) left flank, or (2) right flank. Single crossover events were selected using apramycin, as integration of the pYH7 plasmid conferred apramycin-resistance. (VI) To perform the coding-frame deletion, a second crossing over event was necessary. A coding-frame gene deletion occurred if the second crossing over event used the flanking region not used in the first crossover. If the same flanking region of the pYH7 plasmid used for the first was also used for the second, the strain reverted to wild type. Apramycin sensitivity was restored following the second crossover and detected using replica plating. Figure adapted from[210].

Using this method *S.* sp NRRL 11266 Δtmn8 (**Figure 4.7**), *A. verrucosispora* Δmad10 (**Figure 4.8a**), *A. verrucosispora* Δmad31 (**Figure 4.8b**), *S. longisporoflavus* Δtsn11 (**Figure 4.9a**), and *S. longisporoflavus* Δtsn15 (**Figure 4.9b**) were all successfully created. They were then separately analysed by HPLC-MS for polyether tetronate production.

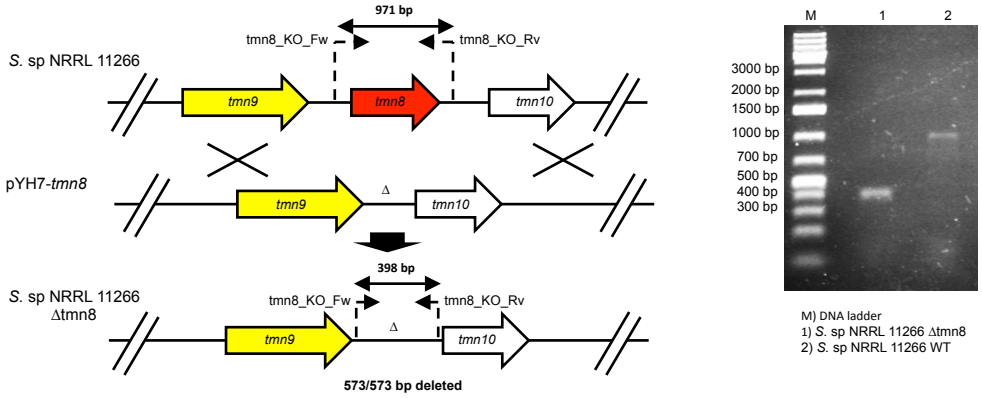


Figure 4.7 Creation of the *S.* sp NRRL 11266 Δtmn8 deletion mutant. Left: Diagram showing the double crossing-over event between pYH7-*tmn8* and the *S.* sp NRRL 11266 chromosome. In total 573/573 bp of *tmn8* were deleted in the Δtmn8 mutant. The primers tmn8_KO_Fw and tmn8_KO_Rv were used to screen the genomic DNA of exconjugants to identify those in which the coding-frame deletion had taken place. PCR product size wild type: 971 bp. PCR product size Δtmn8 mutant: 398 bp. Left: Agarose gel demonstrating the successful creation of the *S.* sp NRRL 11266 Δtmn8 mutant.

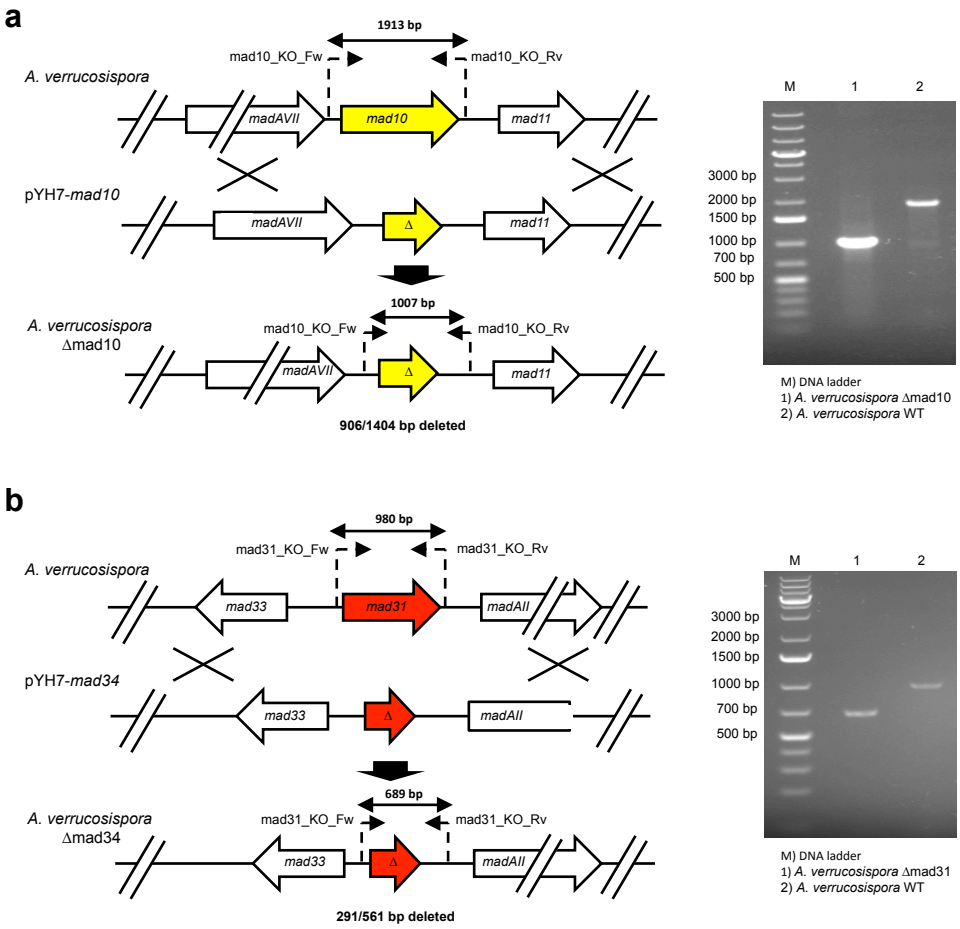


Figure 4.8 Creation of the *A. verrucosispora* Δmad10 and *A. verrucosispora* Δmad31 deletion mutants. **a,** Creation of the *A. verrucosispora* Δmad10. Left: Diagram showing the double crossing-over event between pYH7-*mad10* and the *A. verrucosispora* chromosome. In total 906/1404 bp of *mad10* were deleted in the Δmad10 mutant. The primers mad10_KO_Fw and mad10_KO_Rv were used to screen the genomic DNA of double crossover exconjugants to identify those containing the *mad10* coding-frame deletion. PCR product size wild type: 1913 bp. PCR product size Δmad10 mutant: 1007 bp. Right: Agarose gel demonstrating the successful creation of the *A. verrucosispora* Δmad10 mutant. **b,** Creation of the *A. verrucosispora* Δmad31 mutant. Left: Diagram showing the double crossing-over event between pYH7-*mad31* and the *A. verrucosispora* chromosome. In total 291/561 bp of *mad31* were deleted in the Δmad31 mutant. The primers mad31_KO_Fw and mad31_KO_Rv were used to screen the genomic DNA of double crossover exconjugants to identify those containing the Δmad31 coding-frame deletion. PCR product size wild type: 980 bp. PCR product size Δtsn15 mutant: 689 bp. Right: Agarose gel demonstrating the successful creation of a *S. longisporoflavus* Δmad31 mutant.

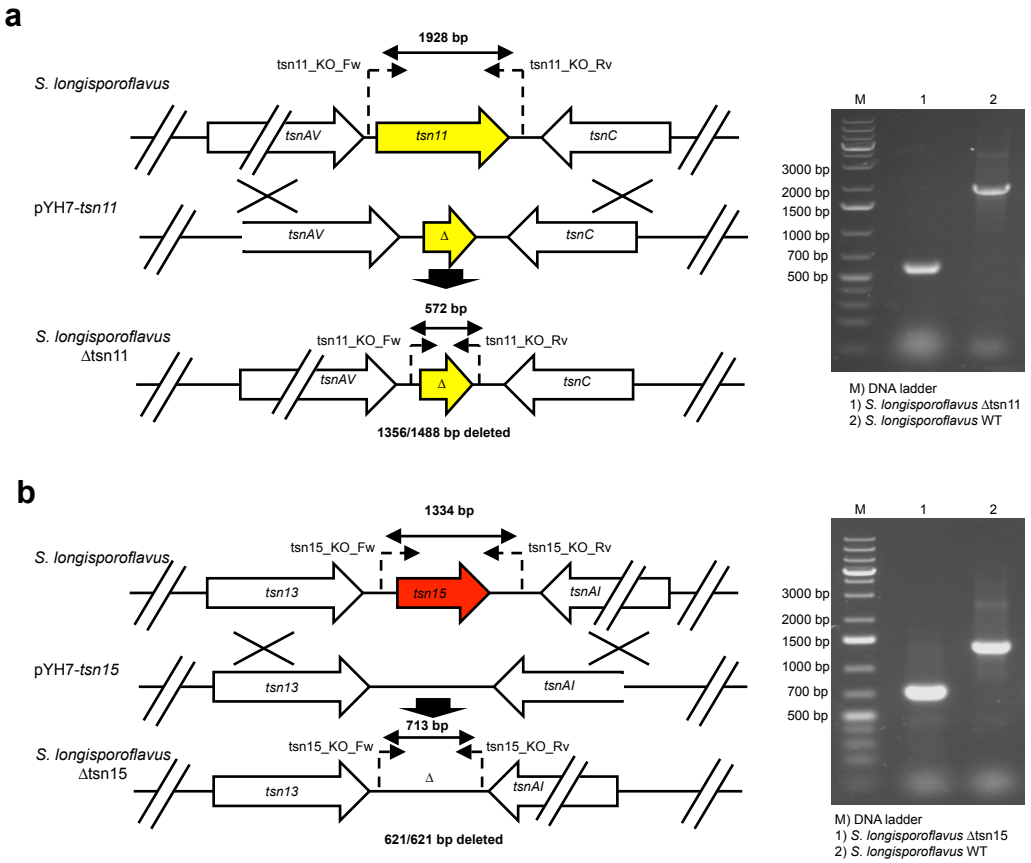


Figure 4.9 Creation of the *S. longisporoflavus* Δtsn11 and *S. longisporoflavus* Δtsn15 deletion mutants. **a,** Creation of *S. longisporoflavus* Δtsn11. Left: Diagram showing the double crossing-over event between pYH7-*tsn11* and the *S. longisporoflavus* chromosome. In total 1356/1488 bp of *tsn11* were deleted in the Δtsn11 mutant. The primers tsn11_KO_Fw and tsn11_KO_Rv were used to screen the genomic DNA of double crossover exconjugants to identify those containing the Δtsn11 coding-frame deletion. PCR product size wild type: 1928 bp. PCR product size Δtsn11 mutant: 572 bp. Right: Agarose gel demonstrating the successful creation of the *S. longisporoflavus* Δtsn11 mutant. **b,** Creation of *S. longisporoflavus* Δtsn15. Left: Diagram showing the double crossing-over event between pYH7-tsn15 and the *S. longisporoflavus* chromosome. In total 621/621 bp of *tsn15* were deleted in the Δtsn15 mutant. The primers tsn15_KO_Fw and tsn15_KO_Rv were used to screen the genomic DNA of double crossover exconjugants to identity those containing the Δtsn15 coding-frame deletion. PCR product size wild type: 1334 bp. PCR product size Δtsn15 mutant: 713 bp. Right: Agarose gel demonstrating the successful creation of the *S. longisporoflavus* Δtsn15 mutant.

4.3 HPLC-MS analysis of polyether tetronate [4+2] cyclase deletion mutants

4.3.1 Analysis of the *S.* sp NRRL 11266 ∆tmn8 deletion mutant

HPLC-MS analysis revealed that tetronomycin production was abolished in *S.* sp. NRRL 11266 ∆tmn8 (**Figure 4.10**), indicating that Tmn8 is essential for tetronomycin biosynthesis. This experiment provided the first evidence that VstJ-like cyclase homologues are essential for polyether tetronate biosynthesis.

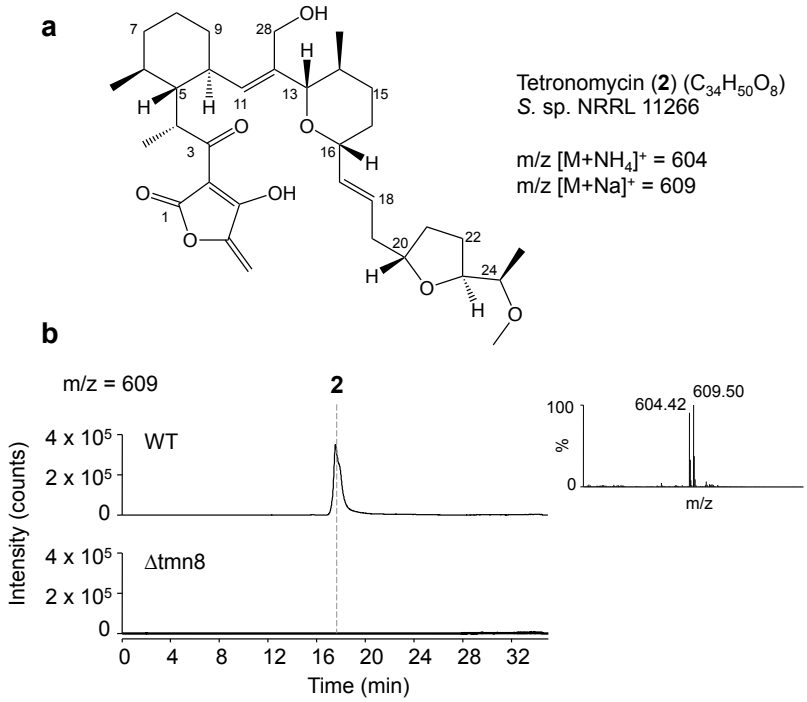


Figure 4.10 HPLC-MS analysis of the *Streptomyces* sp. NRRL 11266 ∆tmn8 mutant. a, The structure and molecular weight of tetronomycin (**2**) produced by *S.* sp. NRRL 11266. **b,** Extracted m/z = 609 chromatograms of *S.* sp. NRRL 11266 wild type (WT) and *S.* sp. NRRL 11266 ∆tmn8. Tetronomycin **2** was detected in *S.* sp. NRRL 11266 WT as ammonium (m/z = 604.42) and sodium (m/z = 609.5) adducts. No tetronomycin production was detected from the *S.* sp. NRRL 11266 ∆tmn8 mutant. Data are representative of three independent experiments.

To verify that the deletion of *tmn8* had not affected the expression of adjacent genes, an attempt was made to complement the ∆tmn8 mutant *in trans.* Such genetic complementation typically utilises a vector capable of integrating into a genomic phage attachment site[210,269]. Once integrated, the desired gene is expressed from the integrated vector. In this study the plasmid pIB139, which integrates into φC31 phage attachment sites, was used[206]. As an example, the pIB139 construct used to deliver *tsn11* in the *S.*

longisporoflavus Δtsn11 mutant is displayed in **Figure 4.11**. Unfortunately in the case of *S.* sp. NRRL 11266 Δtmn8, despite numerous attempts, pIB139-*tmn8* could not be conjugated into *S.* sp. NRRL 11266 Δtmn8. Genetic complementation was therefore not achieved in this case.

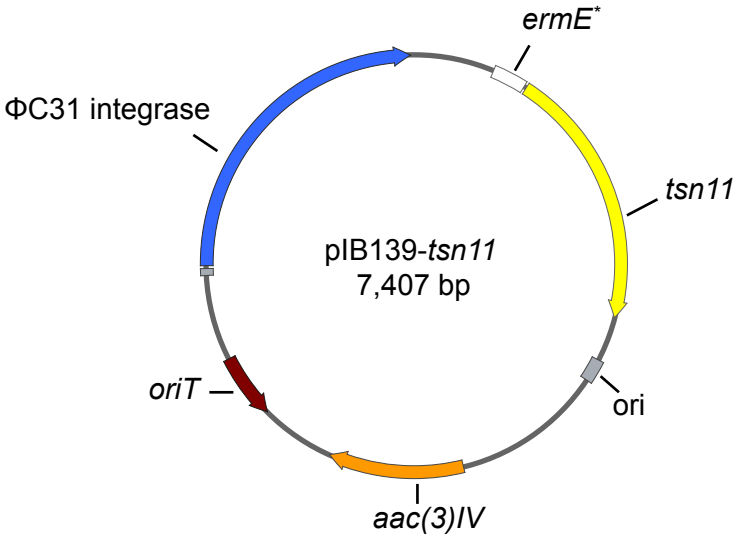


Figure 4.11 Plasmid map of the pIB139-*tsn11*. *tsn11* (yellow) was cloned downstream of the *ermE** promoter (white). Once conjugated into *S. longisporoflavus* Δtsn11, the ΦC31 integrase integrates the plasmid into a chromosomal ΦC31 attachment site. *tsn11* is then expressed from the *ermE** promoter. *acc(3)IV*: apramycin resistance gene. *oriT*: origin of transfer, *ori*: origin of replication.

4.3.2 Analysis of *A. verrucosispora* Δmad10 and *A. verrucosispora* Δmad31 deletion mutants

Tetromadurin production was abolished in the *A. verrucosispora* Δmad10 mutant (**Figure 4.12a**, **Figure 4.12b**), indicating that the PyrE3-like Mad10 is an essential biosynthetic enzyme. Unlike *S.* sp. NRRL 11266 Δtmn8, pIB139-based plasmids could be successfully used for conjugation into *A. verrucosispora*. Introduction of pIB139-*mad10* into *A. verrucosispora* Δmad10 restored tetromadurin production, albeit only to 28% of the wild type level (**Figure 4.12b**). Tetromadurin production in the *A. verrucosispora* Δmad31 mutant was also abolished, producing tetromadurin at just 3% of wild type levels (**Figure 4.12c**). However, *in trans* complementation with *mad31* in *A. verrucosispora* Δmad31 did not restore tetromadurin production (**Figure 4.12c**).

Although *A. verrucosispora* ∆mad10 no longer produced tetromadurin, analysis of the total ion current (TIC) and photodiode array (PDA) spectrum of this mutant revealed that it produced a new metabolite labelled M-17 (**Figure 4.13**). *A. verrucosispora* ∆mad10 produced M-17 at about half the level that *A. verrucosispora* WT produced tetromadurin. M-17 has a different retention time, UV chromophore, and mass spectra from tetromadurin (**Figure 4.13**). The mass spectra of M-17 contains peaks for putative $[M+H]^+$ (m/z = 745), $[M+NH_4]^+$ (m/z = 762), and $[M+Na]^+$ (m/z = 767) ions, all of which were 16 Daltons less than the equivalent ions of tetromadurin (**3**) (**Figure 4.13c**). A mass difference of 16 is diagnostic of an absent hydroxyl group, suggesting M-17 has the molecular formula of $C_{42}H_{64}O_{11}$. Likely candidates for this missing oxygen are either of the C36 or C38 hydroxyl groups. As discussed in Chapter 1, it is likely that the cytochrome P450 enzymes Mad29 and Mad30 add the C36 and C38 hydroxyl groups, respectively (**Figure 4.14**).

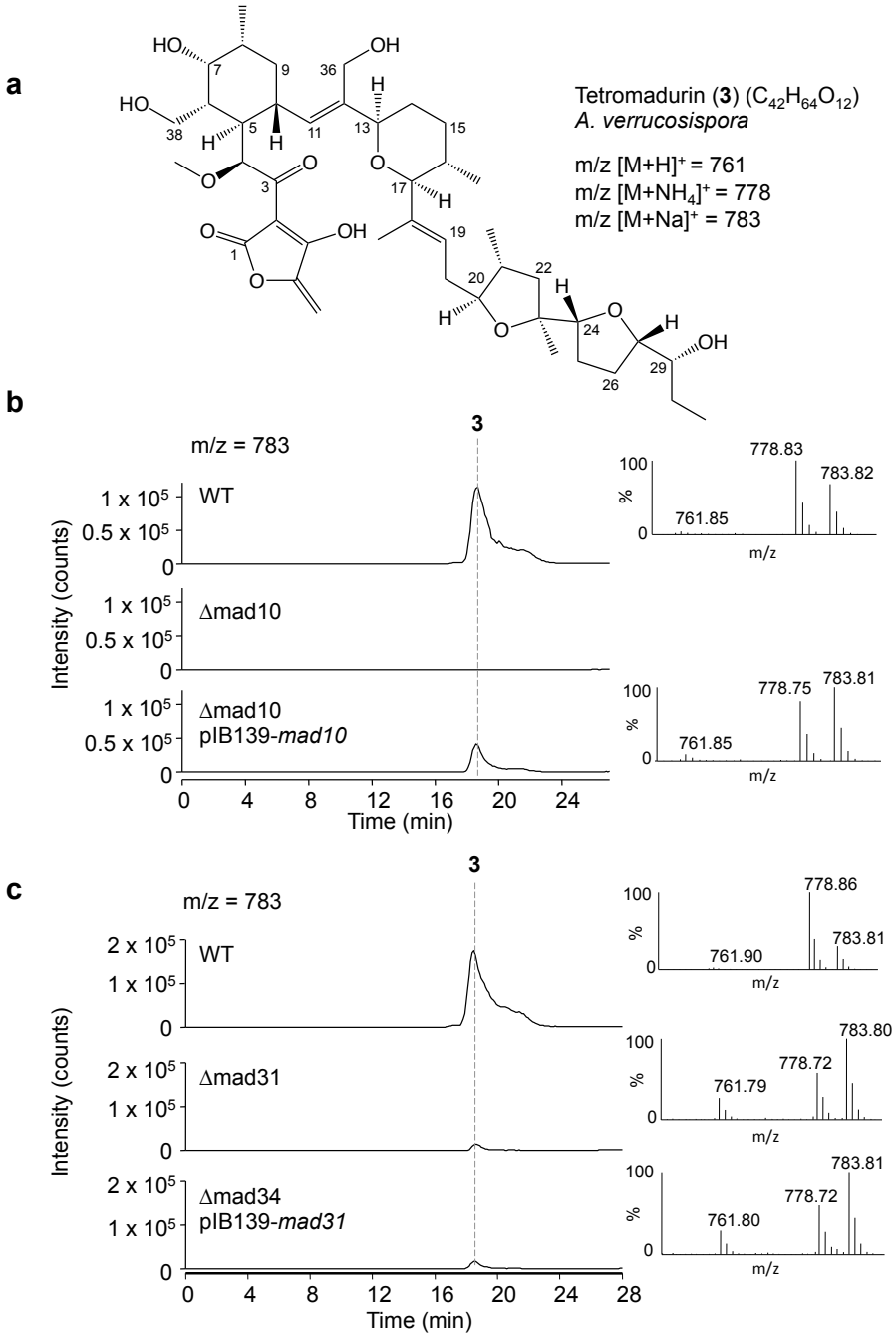


Figure 4.12 HPLC-MS analysis of *A. verrucosispora* ∆mad10 and *A. verrucosispora* ∆mad31. a, The structure and molecular weight of tetromadurin (**3**) produced by *A. verrucosispora*. **b,** Extracted m/z = 783 HPLC-MS spectra from *A. verrucosispora* WT, *A. verrucosispora* ∆mad10, and *A. verrucosispora* ∆mad10 pIB139-*mad10*. **c,** Extracted m/z = 783 HPLC-MS spectra from *A. verrucosispora* WT, *A. verrucosispora* ∆mad31, and *A. verrucosispora* ∆mad31 pIB139-*mad31.* All data are representative of three independent experiments.

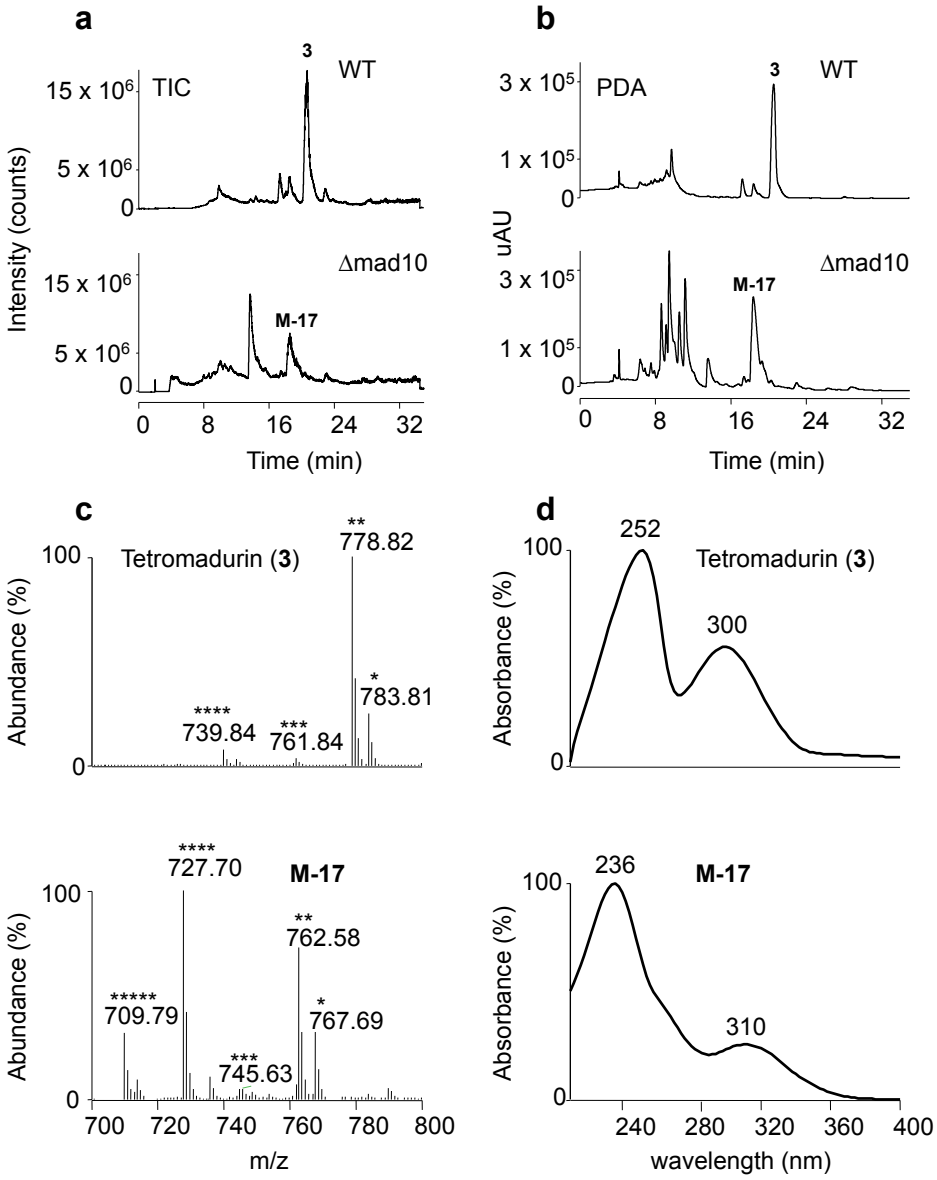


Figure 4.13 Novel tetromadurin analogue produced by *A. verrucosispora* Δmad10. a, Total ion current (TIC) spectra of the crude organic extracts from *A. verrucosispora* WT and *A. verrucosispora* Δmad10. The other peaks represent unrelated species. Tetromadurin (**3**) was produced by *A. verrucosispora* and eluted at 19.8 min. A novel tetromadurin-related metabolite (M-17) was produced by *A. verrucosispora* Δmad10, eluting at 17.4 min. **b,** HPLC-PDA (photodiode array) spectra of the crude organic extracts from *A. verrucosispora* WT and *A. verrucosispora* Δmad10. **c,** Positive ion mode mass spectrum of tetromadurin (**3**) and M-17: *$[M+Na]^+$, **$[M+NH_4]^+$, ***$[M+H]^+$, ****$[M+H-H_2O]^+$, *****$[M+H-2.H_2O]^+$. For tetromadurin (**3**): m/z $[C_{42}H_{64}O_{12}+Na]^+$ = 783.81; m/z $[C_{42}H_{64}O_{12}+NH_4]^+$ = 778.82; m/z $[C_{42}H_{64}O_{12}+H]^+$ = 761.84; m/z $[C_{42}H_{64}O_{12}+H-H_2O]^+$ = 739.84. For M-17: m/z $[C_{42}H_{64}O_{11}+Na]^+$ = 767.69; m/z $[C_{42}H_{64}O_{11}+NH_4]^+$ = 762.58; m/z $[C_{42}H_{64}O_{11}+H]^+$ = 745.63; m/z $[C_{42}H_{64}O_{11}+H-H_2O]^+$ = 727.70; m/z $[C_{42}H_{64}O_{11}+H-H_2O]^+$ = 727.70; m/z $[C_{42}H_{64}O_{11}+H-2.H_2O]^+$ = 709.79. **d,** UV absorption spectra of tetromadurin (**3**) (λ_{MAX} = 252 nm, 300 nm, MeOH) and M-17 (λ_{MAX} = 236 nm, 300 nm, MeOH).

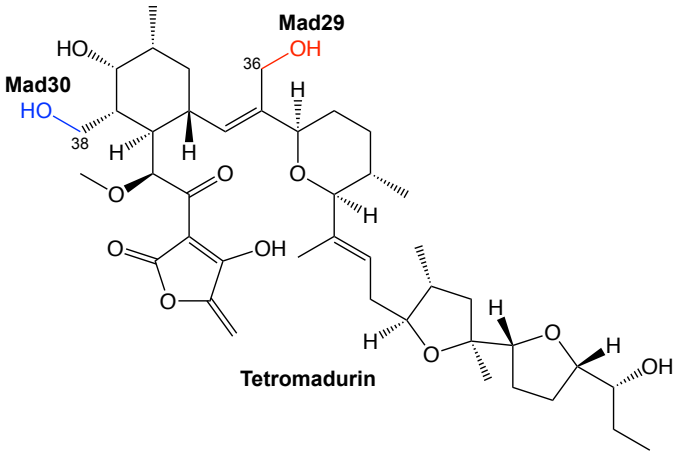


Figure 4.14 Predicted sites of cytochrome P450 hydroxylation on tetromadurin. The OH groups on C36 and C38 (highlighted in red) are likely added by the two cytochrome P450 enzymes, Mad29 and Mad30, respectively, present in the *mad* cluster.

One explanation for these results is that Mad10, rather than Mad29 or Mad30, is responsible for catalysing the hydroxylation at either C36 or C38. However, Mad10 is homologous to FAD-dependent monooxygenases, enzymes that typically catalyse hydroxylations of aromatic compounds[270], exemplified by OxyS in oxytetracycline biosynthesis. In contrast, hydroxylation of C-H bonds is a typical reaction catalysed by cytochrome P450 enzymes[271]. Given the presence of two cytochrome P450 enzymes in the *mad* cluster, it is unlikely that Mad10 is directly responsible for either of these hydroxylations.

An alternative explanation is that Mad10 catalyses a reaction that must precede one of the cytochrome P450-catalysed hydroxylations. Comparison of the M-17 mass spectrum with that of tetromadurin revealed telling differences in the relative abundance of several ions. While the major ions detected for tetromadurin were the $[M+NH_4]^+$ and $[M+Na]^+$ adducts, the base peak for M-17 was the $[M+H-H_2O]^+$ ion (m/z = 727) (**Figure 4.13c**). The $[M+H-H_2O]^+$ ion likely arises through the M-17 $[M+H]^+$ ion losing a hydroxyl group in the mass spectrometer source. An equivalent $[M+H-H_2O]^+$ peak was detectable for tetromadurin (m/z = 739) although only at 12% the abundance of the $[M+NH_4]^+$ base peak. Taken together these data indicate that, in addition to lacking one of the hydroxyl groups added by a cytochrome P450, M-17 also contains a labile hydroxyl group that is readily lost in the mass spectrometer. A possible candidate for the labile hydroxyl group was the C17 hydroxyl, which would only be present if the tetrahydropyran ring had not yet formed. However, the predicted structure would only match the observed m/z values if the

cyclohexane ring had also not formed. It seems reasonable that without the cyclohexane and tetrahydropyran rings one of the cytochrome P450-catalysed hydroxylations, either at C36 or C38, would be unable to occur. From this, two possible structures of M-17 were proposed (**Figure 4.15a**). Both of the proposed intermediate structures could still theoretically form the cyclohexane and tetrahydropyran rings in the ring-closing cascade analogous to that previously proposed for tetronasin and tetronomycin[175,177] (**Figure 4.15b**).

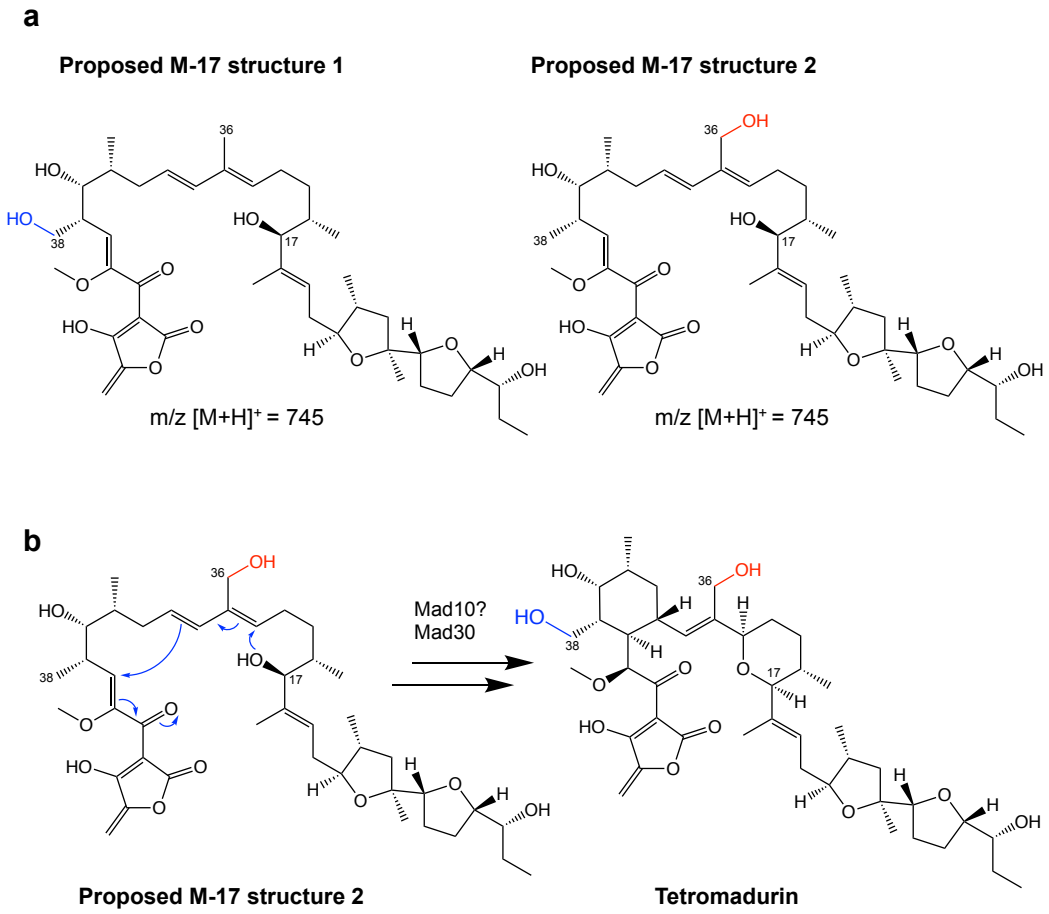


Figure 4.15 Proposed structures of M-17 produced by *A. verrucosispora* Δmad10. a, Potential structures for M-17 that lack the cyclohexane and tetrahydropyran rings and either the C36 (structure 1) or C38 hydroxyl groups (structure 2). **b,** Possible mechanism by which Mad10 could catalyse closure of the tetrahydropyran and cyclohexane, followed by hydroxylation of C38 by Mad30.

If true, this proposal would implicate Mad10 as a cyclase enzyme responsible for closing both the cyclohexane and tetrahydropyran rings of tetromadurin. However, it would not account for why *mad31* is also essential for tetromadurin biosynthesis.

4.3.4 Analysis of the *S. longisporoflavus* Δtsn11 and *S. longisporoflavus* Δtsn15 mutants

Consistent with the results from the *S.* sp NRRL 11266 Δtmn8, *S.* sp NRRL 11266 Δtmn9[111], *A. verrucosispora* Δmad10 and *A. verrucosispora* Δmad31 deletion mutants, tetronasin production was abolished in both the *S. longisporoflavus* Δtsn11 and *S. longisporoflavus* Δtsn15 mutants (**Figure 4.16**). *In trans* complementation of *tsn11* or *tsn15* deletion mutants with the wild type gene restored tetronasin production to wild type levels, confirming the specificity of each gene deletion (**Figure 4.16**). Tsn11 and Tsn15 are therefore essential for tetronasin biosynthesis.

Analogous to the *A. verrucosispora* Δmad10 mutant, *S. longisporoflavus* Δtsn11 produced a new tetronasin analogue. This new metabolite, labelled T-16, had a different retention time, UV chromophore, and MS^3 fragmentation pattern to tetronasin, but the same molecular weight (**Figure 4.17, Appendix Figure 3**). Notably, the UV absorbance spectra of T-16 and M-17 are very similar, each with λ_{MAX} = 236 nm. Also as observed with M-17, the major T-16 ion detected was the $[M+H-H_2O]^+$ ion, whereas for tetronasin the $[M+Na]^+$ was the major ion. T-16 therefore also appears to contain a labile hydroxyl group that is readily lost in the mass spectrometer. The $[M+Na-H_2O]^+$ peak (m/z = 607) did not increase in abundance, indicating that when T-16 was a sodium or ammonium ion adduct, fragmentation of the labile hydroxyl was unfavourable. As suggested for M-17, a possible structure of T-16 that would account for this labile hydroxyl, while retaining the same molecular weight as tetronasin, would be an open form of tetronasin lacking the tetrahydropyran and cyclohexane rings (**Figure 4.18**). To investigate this comprehensive structural studies were planned for the next chapter.

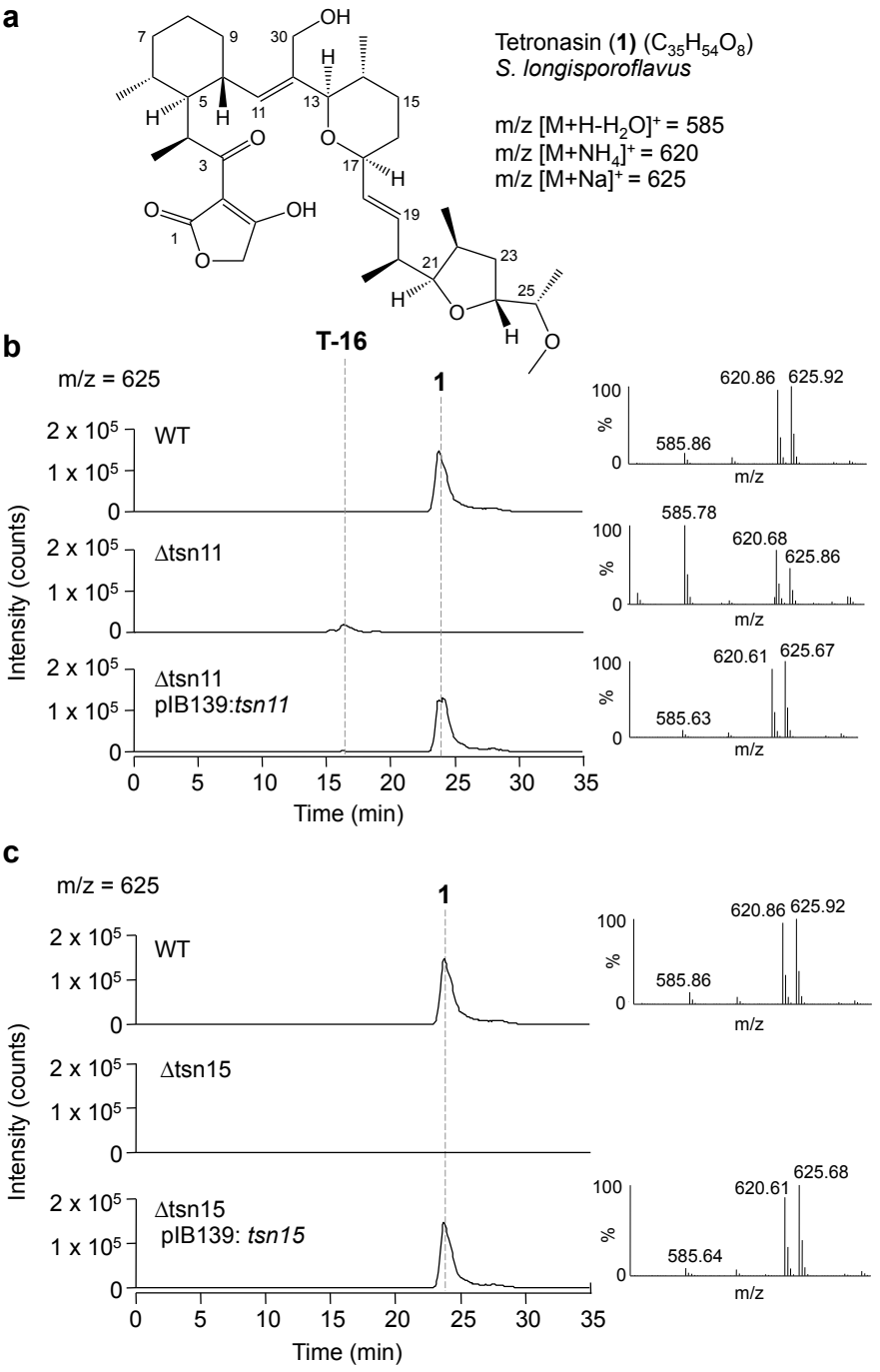


Figure 4.16 HPLC-MS analysis of *S. longisporoflavus* Δtsn11 and *S. longisporoflavus* Δtsn15. a, The structure and molecular weight of tetronasin (**1**) produced by *S. longisporoflavus*. **b,** Extracted m/z = 625 HPLC-MS spectra from *S. longisporoflavus* WT, *S. longisporoflavus* Δtsn11, and *S. longisporoflavus* Δtsn11 pIB139-*tsn11*. **c,** Extracted m/z = 625 HPLC-MS spectra from *S. longisporoflavus* WT, *S. longisporoflavus* Δtsn15, and *S. longisporoflavus* Δtsn15 pIB139-*tsn15*. All data representative of three independent experiments.

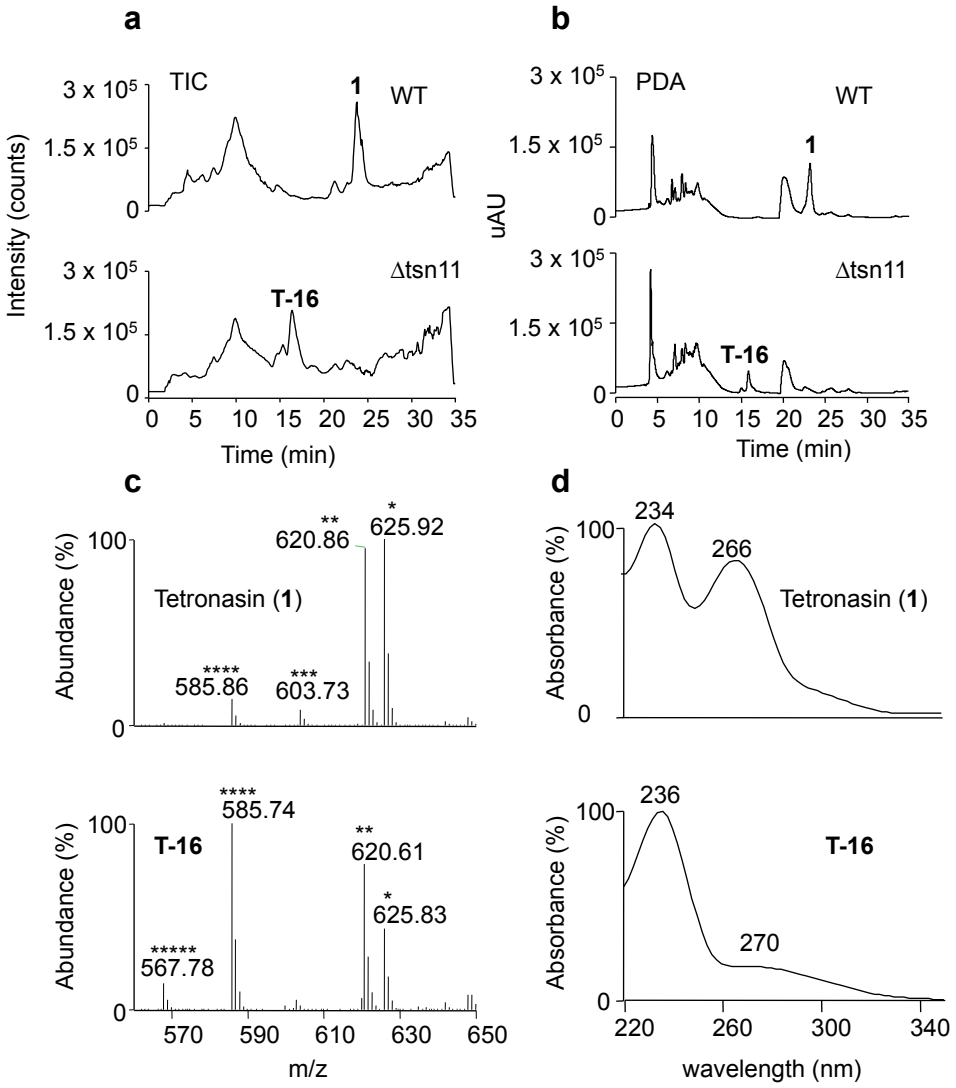


Figure 4.17 Production of a novel tetronasin analogue from *S. longisporoflavus* Δtsn11. a, Total ion current (TIC) spectra of the crude organic extracts from *S. longisporoflavus* WT and *S. longisporoflavus* Δtsn11. Tetronasin (**1**) ($C_{35}H_{54}O_8$) was produced by *S. longisporoflavus* WT, eluting at 24 min. A novel tetronasin analogue (T-16) was produced by *S. longisporoflavus* and eluted at 16.8 min. **b,** HPLC-PDA (photodiode array) spectra of the crude organic extracts from *S. longisporoflavus* Δtsn11 and *S. longisporoflavus* Δtsn15. **c,** Positive ion mode mass spectrum of tetronasin (**1**) and T-16: *$[M+Na]^+$;** $[M+NH_4]^+$; ***$[M+H]^+$; ****$[M+H-H_2O]^+$; *****$[M+H-2.H_2O]^+$. **d,** UV absorption spectra of tetronasin (**1**) (λ_{max} = 234 nm, 266 nm, MeOH) and M-17 (λ_{max} = 236 nm, MeOH).

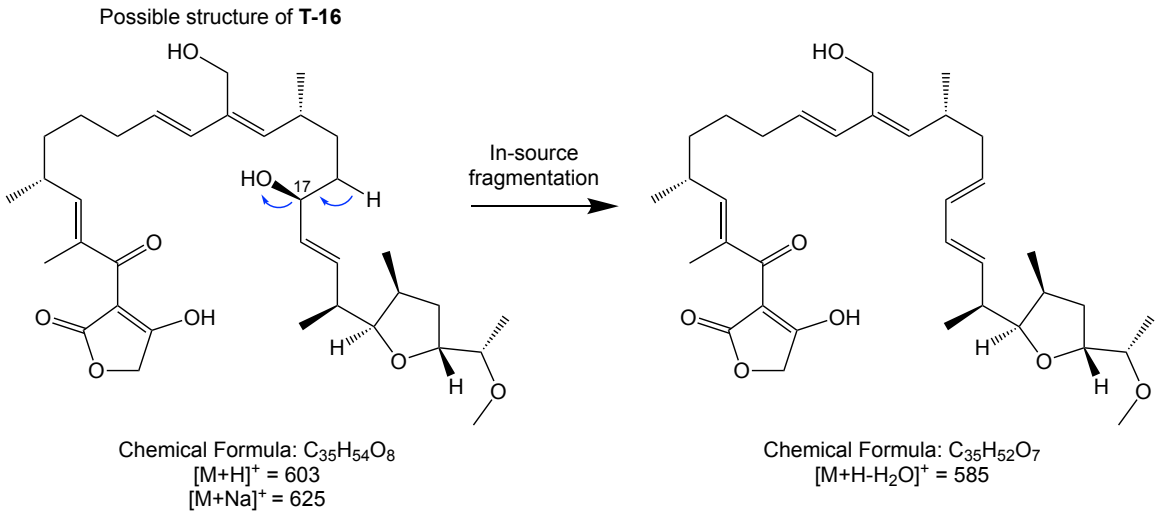


Figure 4.18 Possible structure and in-source fragmentation of T-16. T-16 produced by *S. longisporoflavus* readily fragmented a hydroxyl group in the mass spectrometer, forming $[M+H-H_2O]^+$ as the major ion. One explanation for this could be that the tetrahydropyran and cyclohexane rings of T-16 have not formed, leaving the C17 hydroxyl exposed for in-source fragmentation.

While the *S. longisporoflavus* Δtsn15 mutant no longer produced tetronasin, examination of the TIC spectrum revealed that, in contrast to the *A. verrucosispora* Δmad31 and the *S.* sp NRRL 11266 Δtmn8 mutants, it also produced a new metabolite related to tetronasin. The new metabolite, which was labelled T-21, was produced at a similar level to tetronasin in *S. longisporoflavus* WT strain, and had a different retention time and molecular weight from both T-16 and tetronasin (**Figure 4.19**). Despite being a prominent peak in the TIC spectrum, no corresponding absorbance peak was present in the PxwDA spectrum (**Figure 4.19b**). Intriguingly, the putative $[M+H]^+$, $[M+Na]^+$, $[M+NH_4]^+$, and $[M+H-H_2O]^+$ ions for T-21 were all two mass units greater than the equivalent tetronasin ions (**Figure 4.19c, Figure 4.19d**). Such a mass shift could be caused by the reduction of a double bond or of a keto group, resulting in the addition of two protons. The major ions of T-21, however, were m/z = 538 and m/z = 543, the structures of which could not be confidently predicted. These ions were unique to the *S. longisporoflavus* Δtsn15 mutant and were not detectable in either the WT or in *S. longisporoflavus* Δtsn11. The mass unit difference of 5 between the two major ions suggests that they could reflect ammonium and sodium ion adducts, respectively.

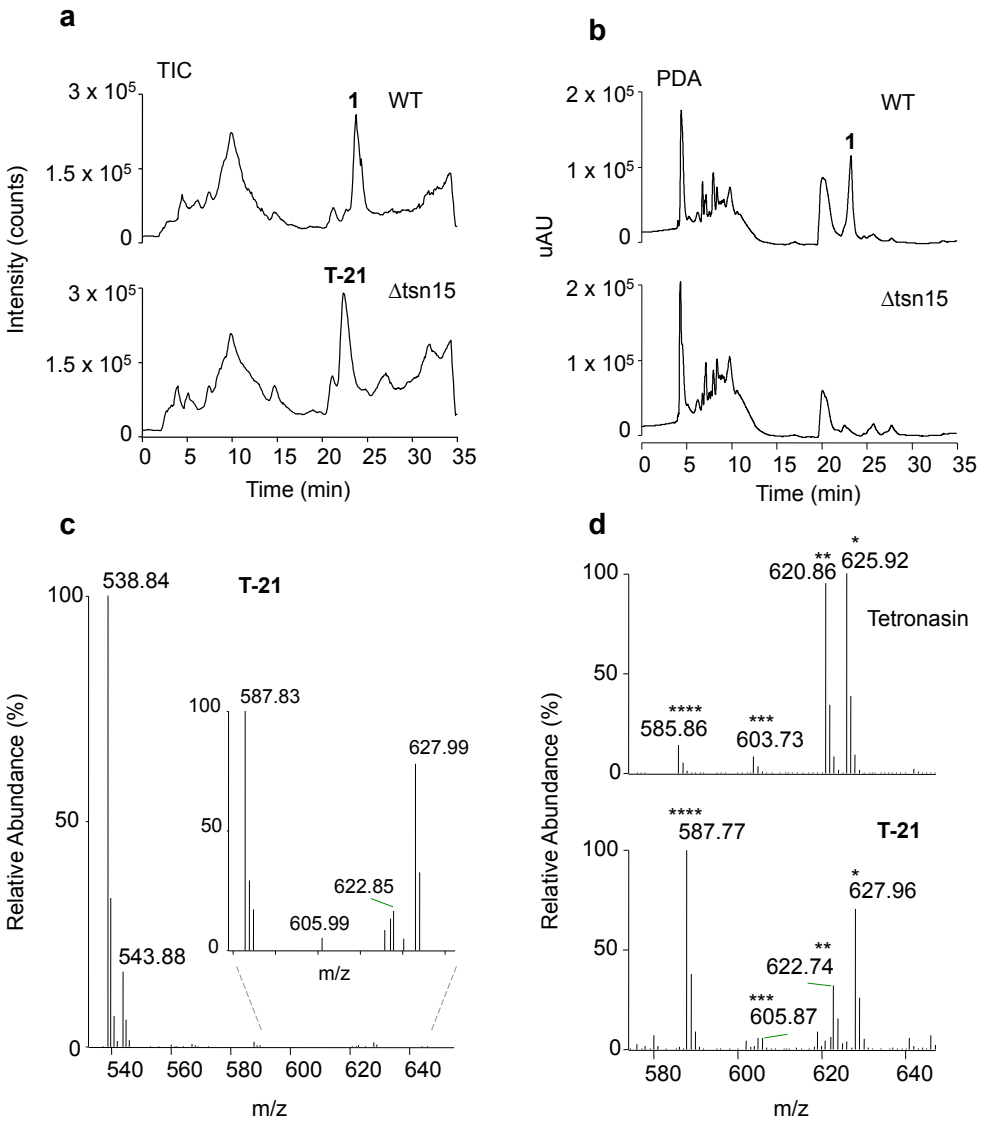


Figure 4.19 Production of T-21 by *S. longisporoflavus* Δtsn15. a, TIC spectra of *S. longisporoflavus* wild type (WT) and *S. longisporoflavus* Δtsn15 showing the production of tetronasin (**1**) and T-21, respectively. **b,** PDA spectra of *S. longisporoflavus* WT and *S. longisporoflavus* Δtsn15. The absorbance peak for tetronasin (**1**) was present in *S. longisporoflavus* WT, but no new peak corresponding to T-21 was present in the *S. longisporoflavus* Δtsn15 spectrum. **c,** Mass spectrum of T-21. **d,** Mass spectra comparison of tetronasin and T-21. *$[C_{35}H_{54}O_8+Na]^+$; **m/z $[C_{35}H_{54}O_8+NH_4]^+$; ***$[C_{35}H_{54}O_8+H]^+$; ****$[C_{35}H_{54}O_8+H-H_2O]^+$.

However, no corresponding $[M+H]^+$ peak at m/z = 520 was detected. It was speculated that the m/z = 538 and m/z = 543 ions were in-source fragmentation products of the $[M+Na]^+$ and $[M+NH_4]$ ions, respectively. However, because the structure of the parent compound was unknown it was difficult to draw a plausible fragmentation mechanism to account for these ions. Further investigation of the identity of T-21 and its relevance to tetronasin biosynthesis is described in Chapter 5 of this thesis.

4.3.5 Creation and analysis of an *S. longisporoflavus* Δtsn11 Δtsn15 double mutant

Given that *S. longisporoflavus* Δtsn11 and the *S. longisporoflavus* Δtsn15 each produced a unique metabolite, reasoned that by creating a *S. longisporoflavus* Δtsn11 Δtsn15 double mutant it would provide insight into the order in which Tsn11 and Tsn15 act. The double mutant was created by deleting *tsn15* from the genome of *S. longisporoflavus* Δtsn11 using the same pYH7 plasmid used for creating *S. longisporoflavus* Δtsn15. The genotype of the mutant was confirmed by PCR amplification of the *tsn11* and *tsn15* loci (**Figure 4.20**).

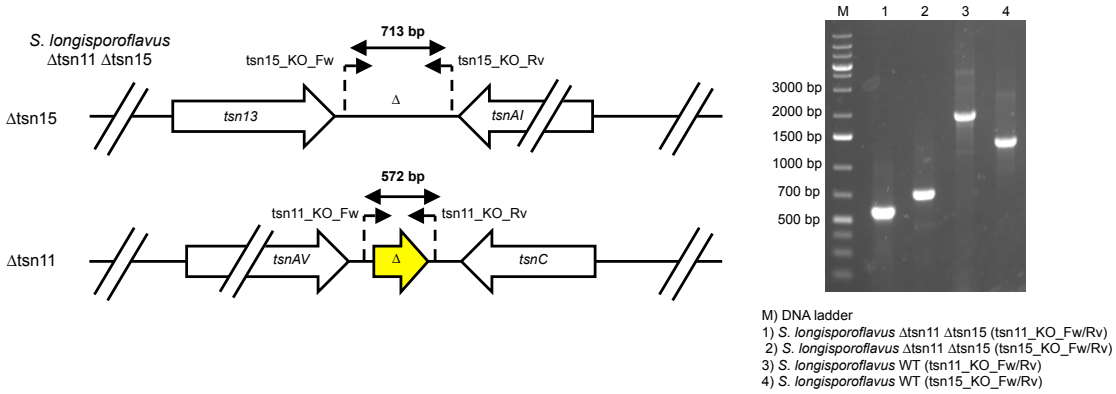


Figure 4.20 Creation of a *S. longisporoflavus* Δtsn11 Δtsn15 double mutant. *tsn15* was deleted from the *S. longisporoflavus* Δtsn11 mutant to create a *S. longisporoflavus* Δtsn11 tsn15 double mutant. The double deletion phenotype was confirmed using PCR and DNA sequencing. Left: the expected Δtsn11 and Δtsn15 PCR product sizes in the *S. longisporoflavus* Δtsn11 Δtsn15 mutant were 713 bp and 572 bp, respectively. Right: agarose gel showing the products of the PCR reactions from the genomic DNA of *S. longisporoflavus* Δtsn11 Δtsn15 and *S. longisporoflavus* WT, confirming the double deletion genotype.

HPLC-MS analysis of *S. longisporoflavus* Δtsn11 Δtsn15 revealed that, like *S. longisporoflavus* Δtsn11, it also produced T-16 (but not T-21) (**Figure 4.21**). These data indicated epistatic silencing of the Δtsn15 phenotype in the Δtsn11 Δtsn15 double mutant, suggesting that Tsn11 acts before Tsn15.

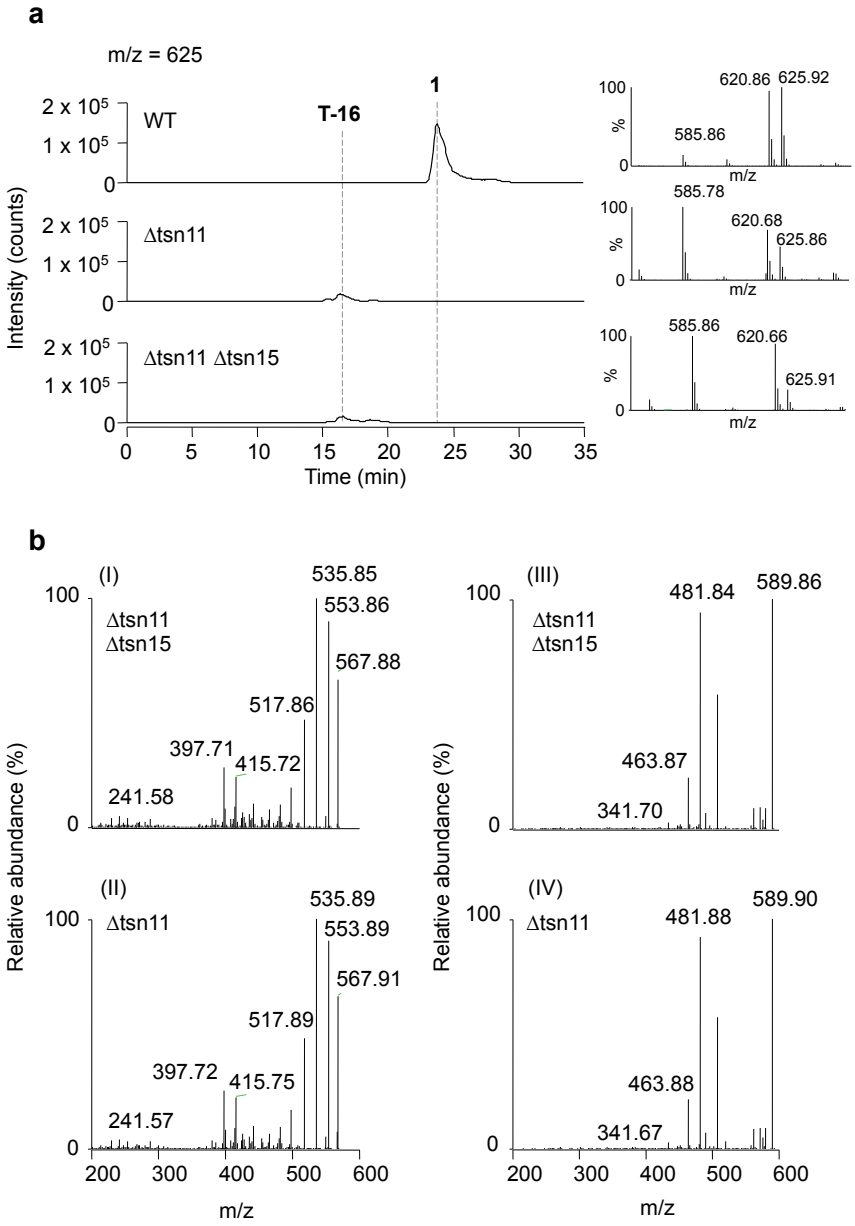


Figure 4.21 Production of T-16 by *S. longisporoflavus* Δtsn11 Δtsn15. a, Extracted m/z = 625 HPLC-MS of *S. longisporoflavus* WT, *S. longisporoflavus* Δtsn11, and *S. longisporoflavus* Δtsn11 Δtsn15. *S. longisporoflavus* WT produced tetronasin while *S. longisporoflavus* Δtsn11 and *S. longisporoflavus* Δtsn11 Δtsn15 both produced T-16. **b,** The production of T-16 by both *S. longisporoflavus* Δtsn11 and *S. longisporoflavus* Δtsn11 Δtsn15 was confirmed by comparing the m/z = 603 $[M+H]^+$ MS^3 of the T-16 produced by (I) *S. longisporoflavus* Δtsn11 Δtsn15, with (II) *S. longisporoflavus* Δtsn11, and the m/z = 625 $[M+Na]^+$ MS^3 of T-16 produced by (III) *S. longisporoflavus* Δtsn11 Δtsn15 and (IV) *S. longisporoflavus* Δtsn11. Data are representative of three independent experiments.

4.3 Chapter 4 Discussion

The results from this chapter provided direct experimental evidence that the *pyrE3* and *vstJ* [4+2] cyclase homologues in the *tmn, tsn,* and *mad* clusters are essential for polyether tetronate biosynthesis. When the *pyrE3* homologues were investigated, the *S.* sp NRRL 11266 ∆tmn9 mutant produced no polyether-related metabolite as previously reported[111] but the *S. longisporoflavus* ∆tsn11 and *A. verrucosispora* ∆mad10 mutants did each produce a new metabolite: T-16 and M-17, respectively. The mass and UV/Vis spectra of T-16 and M-17 suggested that they are analogous compounds, potentially lacking both the tetrahydropyran and the cyclohexane ring. M-17 also appeared to lack one of the hydroxyl groups added by a cytochrome P450 enzyme, but otherwise the MS evidence indicated that it contained the tetronate ring, tetrahydrofuran rings, and that one hydroxylation had taken place. T-16 had the same molecular weight as tetronasin, indicating that tetronate formation, tetrahydrofuran formation, C25 O-methylation, and C30 hydroxylation had all occurred and are not dependent on Tsn11. A notable difference between tetronasin and tetromadurin is that tetronasin only has one cytochrome P450-catalysed hydroxylation (at C30) while tetromadurin has two (at C36 and C38). Assuming that the equivalent carbon to tetronasin C30 is hydroxylated in M-17, C36, it would suggest that the missing hydroxyl group in M-17 is that on C38 (as shown in **Figure 4.15b**).

Deletion of the *vstJ*-like genes *tmn8, tsn15,* and *mad31* in their respective BGCs provided the first functional evidence that these [4+2] cyclase homologues are also essential for polyether tetronate biosynthesis. A new metabolite was produced by the *S. longisporoflavus* ∆tsn15 mutant (T-21), but not by the *S.* sp NRRL 11266 ∆tmn8 or *A. verrucosispora* ∆mad31 mutant. Why some deletion mutants produce novel metabolites and others do not is unknown. However, *S. longisporoflavus* 8E36, is an industrial producer of tetronasin reported to produce up to 500 mg/L[209]. The high tetronasin titre of this strain could be a factor in why the *S. longisporoflavus* ∆tsn11 and *S. longisporoflavus* ∆tsn15 mutants produced detectable polyether tetronate analogues, while the other strains did not.

Production of M-17 and T-16 compared to that of tetromadurin and tetronasin, respectively, was diminished. *A. verrucosispora* ∆mad10 produced M-17 at about half the level at which *A. verrucosispora* WT produced tetromadurin. Likewise, *S. longisporoflavus* ∆tsn11 produced T-16 at ca. 30% the level *S. longisporoflavus* WT produced tetronasin. Several

factors could explain the decrease in production level of these metabolites. The first is that M-17 and T-16 may be unstable in cellular conditions. The second is that Tsn11/Mad10 normally act while the polyketide intermediate is still attached to the PKS, and that without these enzymes subsequent chain extensions are less efficient. The timing of Tsn11 and Tsn15 action is investigated further in **Chapter 5**.

The genetic complementation experiments described in this chapter had mixed success. Using pIB139 to introduce *tsn11* into *S. longisporoflavus* ∆tsn11 and *tsn15* into *S. longisporoflavus* ∆tsn15 both restored tetronasin production to wild type levels (**Figure 4.16**). The *A. verrucosispora* complementation experiments, on the other hand, were less successful. Genetic complementation of *mad10* into *A. verrucosispora* ∆mad10 restored production of tetromadurin to 30% of wild type levels, whereas genetic complementation of *A. verrucosispora* ∆mad31 with *mad31* into did not rescue production at all (**Figure 4.12**). It cannot be ruled out that the failure to rescue the *A. verrucosispora* ∆mad31 phenotype was due to polar effects of the ∆mad31 deletion on adjacent genes. However, given that *mad31* was deleted in its coding frame, such polar effects are unlikely[210]. It is more likely that the pIB139 constructs are not expressed well in the rare actinomycete *A. verrucosispora*. Expression from constitutive promoters such as *ermE** (as found in pIB139) are known to be variable in different conditions/hosts[272] and the position of the integrated gene in the chromosome can also profoundly affect its expression[273]. Poor expression could explain why only partial rescue of tetromadurin production was obtained in the *A. verrucosispora* ∆mad10 pIB139-*mad10* mutant. Many alternative promoter systems are available for such complementation experiments[272]. Future experiments could therefore try to find alternative promoters optimised for the heterologous expression of genes in *A. verrucosispora*.

The final experiment performed in this chapter involved creating an *S. longisporoflavus* ∆tsn11 ∆tsn15 double mutant. Analysis of this double mutant revealed that it produced T-16, like the *S. longisporoflavus* ∆tsn11 mutant. This result strongly suggested that Tsn11 and Tsn15 act sequentially in that order. It is notable that the predicted order of Tsn11 (PyrE3 homologue) and Tsn15 (PyrI4 homologue) is the same order as the homologous proteins act in spirotetronate/tetramate pathways[161]. However, it is highly unlikely that Tsn11 and Tsn15 catalyse [4+2] cycloadditions patterned on PyrE3 and PyrI4.

The mass spectra of T-16 and M-17 gave encouraging evidence that Tsn11 and Mad10 might indeed be cyclases responsible for the biosynthesis of the cyclohexane and tetrahydropyran rings. However, to support this hypothesis it was necessary to purify T-16 and M-17 and characterise their structure. As *S. longisporoflavus* was easier to manipulate genetically and ferment than *A. verrucosispora,* the next chapter focuses on the structural characterisation of T-16 produced by *S. longisporoflavus* Δtsn11.

Chapter 5: *in vitro* reconstitution of tetronasin cyclohexane and tetrahydropyran biosynthesis

The findings of the previous chapter established that specific deletion mutants of *S. longisporoflavus* and *A. verrucosispora* lacking putative [4+2] cyclase genes accumulate novel tetronate-related metabolites apparently lacking the central rings of the mature antibiotics. This opened the exciting possibility of being able not only to isolate and determine the structure of these metabolites, but also to test their competence as substrates for each of the candidate cyclases expressed as recombinant proteins from *E. coli*, aiming essentially to reconstitute formation of the central rings *in vitro*. Success in this would not only define the role of each enzyme in this novel transformation, but also open the way to a detailed structural and mechanistic exploration of these catalysts.

Some sections of this chapter were done in collaboration with Dr. Finian Leeper (University of Cambridge, UK) and Robert Jenkins and Dr. Manuela Tosin (University of Warwick, UK). All contributions will be explicitly stated.

5.1 The structure of T-16

5.1.1 Fermentation of *S. longisporoflavus* Δtsn11

In the previous chapter, small-scale cultures of *S. longisporoflavus* Δtsn11 were grown to produce analytical amounts of T-16 for HPLC-MS analysis. To determine its structure by NMR, however, it was necessary to significantly increase the amount and purity of T-16. Collecting good NMR data typically requires at least 1 mg of highly purified compound, but the more the better, especially in case of large and complex natural products.

S. longisporoflavus Δtsn11 was found to produce T-16 at ca. 30% of the level at which *S. longisporoflavus* WT produced tetronasin. Under optimal conditions, *S. longisporoflavus* 8E36 has been reported to produce tetronasin at levels as high as 500 mg/L[209]. *S. longisporoflavus* Δtsn11 might therefore produce up to 150 mg/L of T-16, suggesting a

relatively modest scale-up would be sufficient. Unfortunately, although T-16 was reliably produced by *S. longisporoflavus* Δtsn11 in 50 mL and 100 mL liquid cultures, production levels dropped to trace amounts when scaled up to 400 mL cultures in 2 L Erlenmeyer flasks. In these larger volumes the growth of *S. longisporoflavus* Δtsn11 was also clearly suboptimal. In contrast, growth of *S. longisporoflavus* Δtsn11 on solid agar reliably produced T-16 and the metabolite could be efficiently extracted using minimal amounts of ethyl acetate. For the initial scale-up, 100 12 x 12 cm tsn-medium-B agar plates were inoculated with *S. longisporoflavus* Δtsn11 and grown at 30 °C for seven days (**Figure 5.1**). Each plate was then extracted three times using ethyl acetate to yield, after removal of the solvent, a total of 7 g of crude organic extract.

Figure 5.1 Growth of *S. longisporoflavus* Δtsn11 on tsn-medium-B agar plates. On tsn-medium-B *S. longisporoflavus* Δtsn11 grew as a dark brown/purple mycelia culture with a grape-like odour.

5.1.2 Purification of T-16

T-16 was isolated from the crude organic extract using reverse-phase chromatography followed by semi-preparative HPLC. The *S. longisporoflavus* Δtsn11 crude organic extract was redissolved in methanol and loaded onto C18 solid-phase extraction cartridges (**Figure 5.2**).

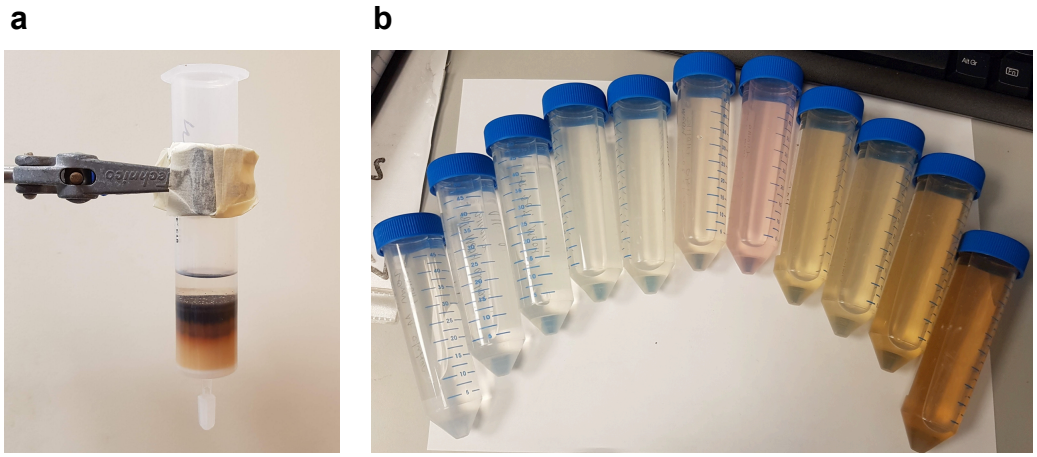


Figure 5.2 C18 solid-phase extraction cartridge used for the purification of T-16. a, The *S. longisporoflavus* Δtsn11 crude organic extract was redissolved in methanol and loaded onto C18 solid phase extraction cartridges. Non-polar species (dark brown fraction) remained at the top of the C18 resin bed, while more polar species, including T-16, could be eluted using 65% methanol. **b,** Example of fractions collected from the C18 solid-phase extraction cartridge (left to right increasing methanol concentration).

Methanol concentrations ranging from 10-100% were passed through the column and the flow-through was collected in fractions. The different fractions were analysed by HPLC or HPLC-MS to identify those containing T-16. From these initial optimisation experiments a reliable isocratic elution protocol was developed for the initial purification of T-16. Passing 65% methanol through the column eluted T-16 after the majority of more polar species had already eluted, but before less polar species started to elute. Typical HPLC traces showing the elution of T-16 across six 65% methanol elution fractions are displayed in **Figure 5.3**.

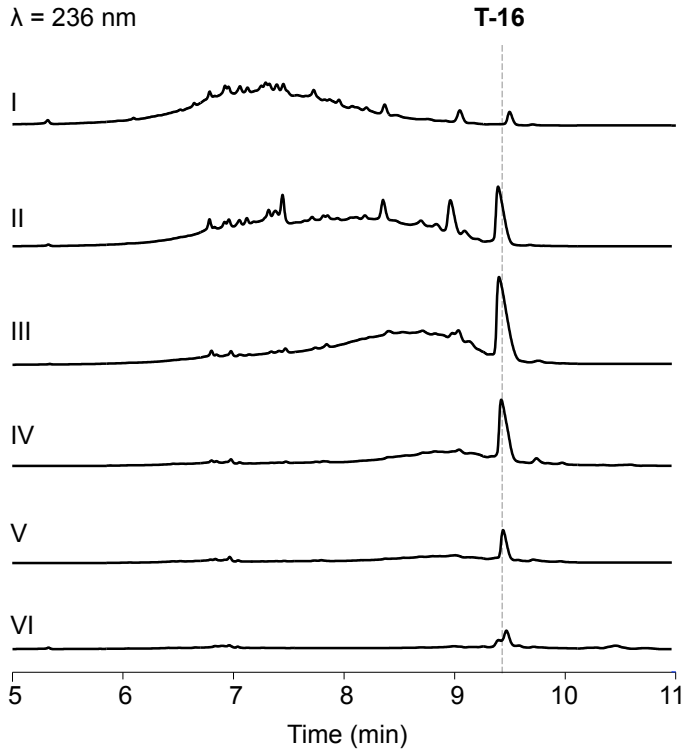


Figure 5.3 HPLC traces from the purification of T-16 using C18 solid phase extraction cartridges. The organic extract from *S. longisporoflavus* Δtsn11 was loaded on C18 solid-phase extraction cartridges. Isocratic elution of T-16 was performed using 65% methanol. Following addition of the 65% methanol to the column, six different election fractions were collected one after the other (I-VI). Each of the fractions was analysed by HPLC for the presence of T-16. All of the fractions in which T-16 was the major component (III-V) were pooled together for the next stage of purification.

The elution fractions in which T-16 was the major component were pooled for further purification. The second round of purification was performed using the fraction collector of a semi-preparative HPLC instrument. A representative spectrum from the semi-preparative HPLC showing the collection of T-16 is displayed in **Figure 5.4**.

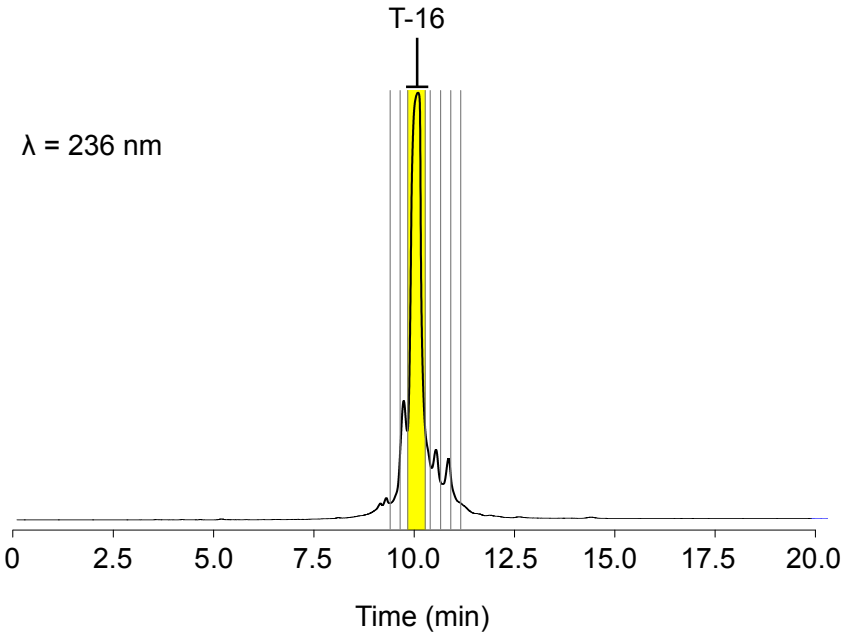
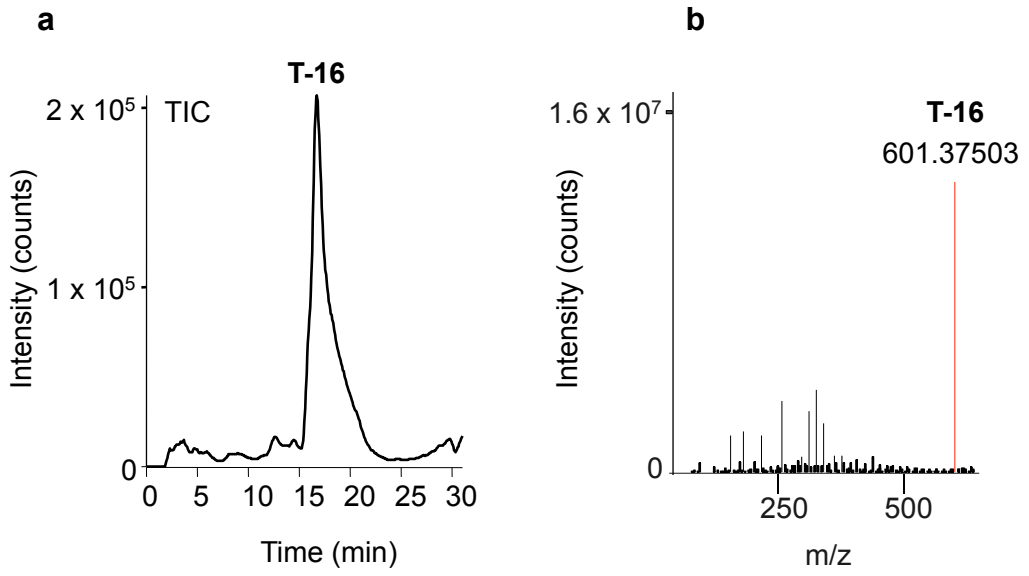


Figure 5.4 The purification of T-16 using semi-preparative HPLC. The figure shows a representative example of the fraction collection program used to purify T-16. The grey lines show the boundaries of the different fractions automatically collected by the instrument. T-16 was collected in the third fraction (highlighted in yellow). The other peaks represent unrelated species.

In this manner a total of 12.5 mg of T-16 was purified from the 100 tsn-medium-B agar plates. The purity of T-16 was assessed by HPLC-MS (**Figure 5.5a**) and high-resolution mass spectrometry confirmed that the molecular formula of T-16 was indeed the same as tetronasin ($C_{35}H_{54}O_8$) (**Figure 5.5b, Figure 5.5c**).



c

Name	Formula	Neutral Mass	Observed neutral mass	Observed m/z	Adduct	Error (ppm)
T-16	$C_{35}H_{54}O_8$	602.38187	602.3824	601.3751	−H	0.8

Figure 5.5 Purification and high-resolution mass spectrometry of T-16. **a,** Total ion current HPLC chromatogram of purified T-16. **b,** High resolution mass spectrum of T-16 obtained on negative ion mode. The T-16 peak is highlighted in red. **c,** Table displaying the data obtained from the high-resolution mass of spectrum of T-16, confirming that the molecular formula was $C_{35}H_{54}O_8$.

5.1.3 NMR of T-16

Prior to analysis by NMR, mass spectrometry was used to confirm that T-16 had an additional hydroxyl group compared to tetronasin. To achieve this, purified T-16 was mixed in a solution of deuterated methanol and D_2O to substitute all exchangeable protons, i.e., those of hydroxyl groups, with deuterium before being directly injected into the mass spectrometer. As predicted in Chapter 4, T-16 was found to have three exchangeable protons, in contrast to the two of tetronasin (**Figure 5.6**)[268].

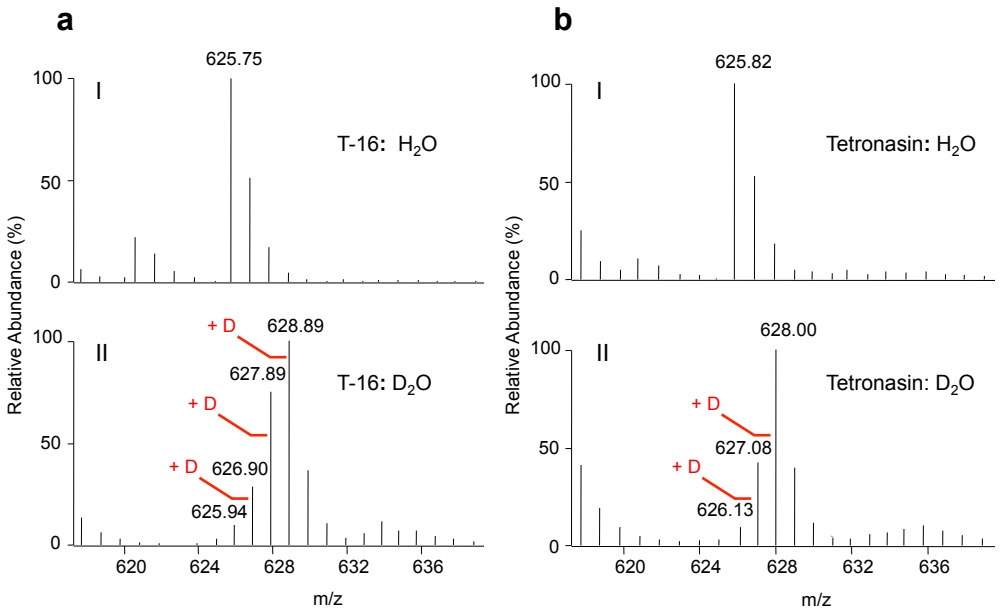


Figure 5.6 Deuterium labelling of exchangeable protons in T-16 and tetronasin. a, T-16 was directly injected into the mass spectrometer after being mixed with (I) methanol and H_2O or (II) deuterated methanol and D_2O. T-16 gained three mass units after being mixed with deuterated methanol and D_2O, indicating three exchangeable protons (hydroxyl groups). **b,** Tetronasin control (I) A direct injection of tetronasin after mixing with methanol H_2O. (I) A direct injection of tetronasin after mixing with deuterated methanol and D_2O, demonstrating that tetronasin had two exchangeable protons (hydroxyl groups)[268].

Purified T-16 was then submitted for extensive NMR characterisation consisting of ^{1}H, ^{13}C, COSY, HSQC, HMBC, and NOESY experiments. The NMR spectra were analysed in collaboration with Dr. Finian Leeper (University of Cambridge). From the NMR data the structure of T-16 could be confidently determined (**Figure 5.7, Appendix Figure 10-15, Appendix Table 1**).

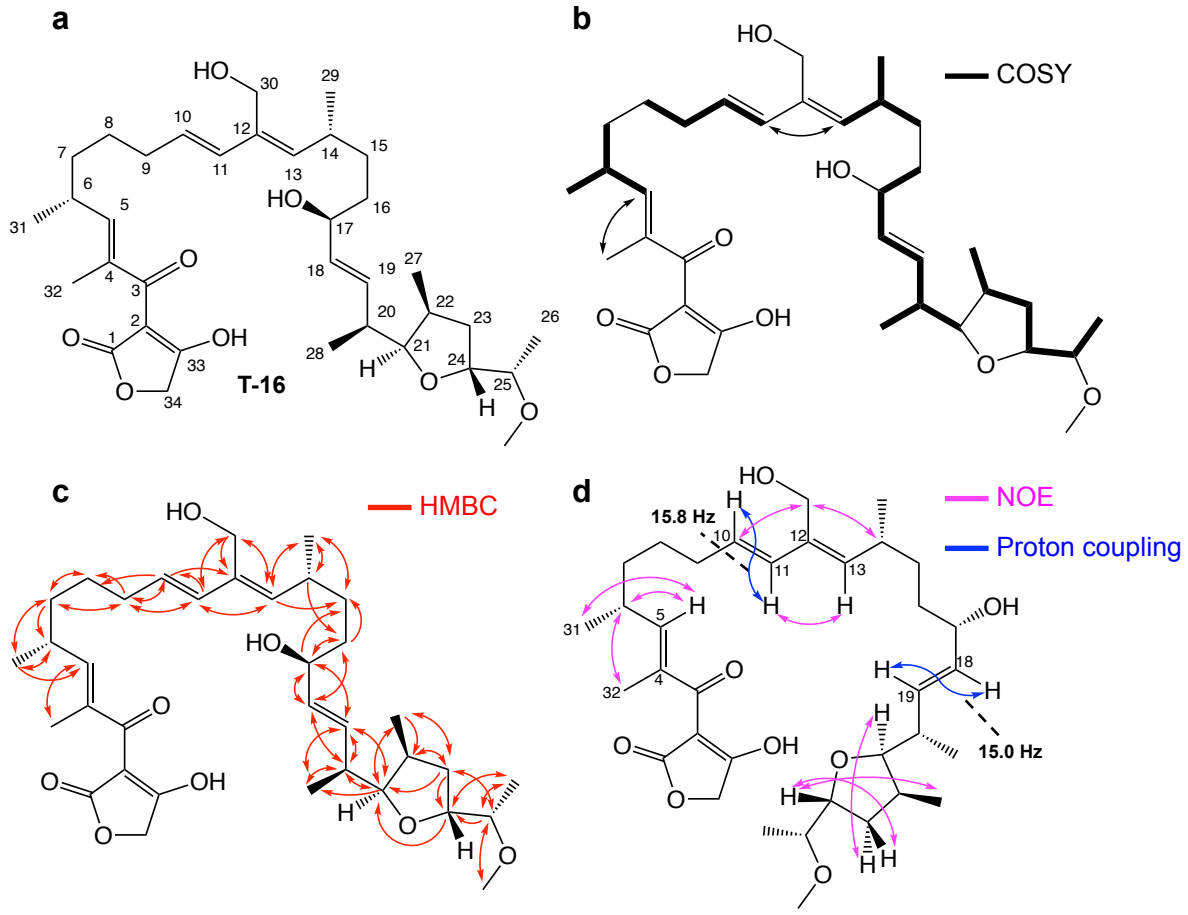


Figure 5.7 The structure of T-16 as determined by NMR spectroscopy. a, The structure of T-16 with atom labels. **b,** ^{1}H-^{1}H COSY correlations for T-16. **c,** ^{1}H-^{13}C HMBC correlations of T-16. Arrows point from proton to carbon. **d,** Selected ^{1}H-^{1}H NOESY correlations and ^{1}H coupling constants of T-16.

As proposed in Chapter 4, T-16 is an open form of tetronasin lacking the cyclohexane and tetrahydropyran rings, but containing the tetronate and tetrahydrofuran rings, the C25 methoxy group, and the C30 hydroxylation. Although the HRMS of T-16 confirmed its molecular formula was identical to that of tetronasin ($C_{35}H_{54}O_8$), no signals corresponding to C3 or the tetronate ring (C1, C2, C33, C34) were present in the ^{13}C spectrum. The absence of these signals has been reported previously for simple 3-acyl tetronates[274] and is due to rotation about the C2-C3 bond being intermediate in speed on the NMR time-scale, causing extreme line broadening of the ^{13}C signals. Signals for these carbon atoms have been detected for tetronasin[174,181], probably because its rotational flexibility is reduced by coordination to a sodium atom. T-16 is therefore clearly less effective at binding a sodium ion than tetronasin. The rest of the ^{1}H and ^{13}C signals could be readily assigned to atoms in T-16. The COSY and HMBC signals from C-17 and C28, including the methoxy group, were very similar to those previously measured for tetronasin[174,181]. The ^{1}H

spectrum of T-16 revealed six olefinic protons, whereas tetronasin only had three[174,181]. In the case of tetronasin, the olefin peaks correspond to protons at the C11-C12 and C18-C19 double bonds. In contrast, T-16 has double bonds at C4-C5, C10-C11, C12-C13, and C18-C19 (the exact locations predicted by the TsnAI-TsnAV enzymes – see **Figure 1.44**). The ^{1}H coupling constants of H10-H11 and H18-H19 double bonds were 15.8 Hz and 15.0 Hz, respectively, indicating that the double bonds are *trans*. The configurations of the double bonds with only a single attached proton (C4-C5 and C12-C13) were determined by the NOESY spectrum as also being *trans*. The C4-C5 double bond was determined as *trans* due to the methyl protons of C32 having NOEs to the H6, but not to the H5 proton. The C12-C13 double bond was likewise determined as *trans* due to an NOE between the H13 proton and H11 proton, but not between the H13 and H30 protons (as would have been expected for a *cis* configuration). The NOESY also revealed that the relative stereochemistry of the tetrahydrofuran ring was the same as in tetronasin. The relative stereochemistries of the other chiral centres (C6, C14, C17, C20, and C25) could not be assigned from the data, but are unlikely to differ from tetronasin.

That T-16 did not have the cyclohexane and tetrahydropyran rings provided the strongest evidence to date that formation of these rings is enzyme catalysed, possibly by Tsn11. The structure of T-16 also demonstrated that tetrahydrofuran formation, C25 O-methylation, and C30 hydroxylation are not dependent on cyclohexane and tetrahydropyran ring formation. However, whether T-16 was a real tetronasin intermediate or a metabolic shunt product was unknown. To investigate this, it was essential to test whether or not T-16 could be converted into tetronasin *in vitro* using purified Tsn11. Analysis of the *S. longisporoflavus* Δtsn11 Δtsn15 double mutant (**Figure 4.21**) had suggested that Tsn15 acts after Tsn11, indicating that both enzymes might be required for cyclohexane and tetrahydropyran formation. Protein purification protocols were therefore developed for both Tsn11 and Tsn15.

5.2 in vitro reconstitution of ring formation

5.2.1 Purification of recombinant Tsn11 and Tsn15

The *tsn11* and *tsn15* genes were amplified from the genomic DNA of *S. longisporoflavus* and individually cloned into the *E. coli* expression vector pET28a(+). The pET28a(+) vector added an *N*-terminus polyhistidine tag in-frame with each of the genes, enabling Tsn11 and Tsn15 to be purified using nickel-affinity chromatography (**Figure 5.8**). The identity of each purified protein was confirmed using mass spectrometry (**Figure 5.9**).

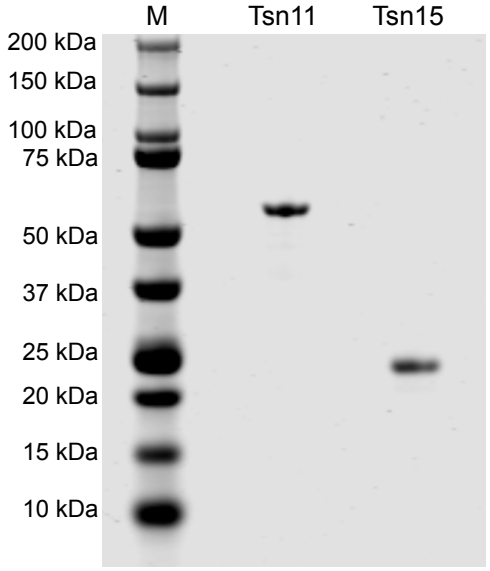


Figure 5.8 SDS-PAGE gel of the purified Tsn11 and Tsn15. Tsn11 and Tsn15 were individually purified using nickel-affinity chromatography and analysed for purity on an SDS-PAGE gel. On the SDS-PAGE gel from left to right: Marker (M); Tsn11, 55.5 kDa; Tsn15, 24.6 kDa. Protein bands were visualised using Coomassie Blue stain.

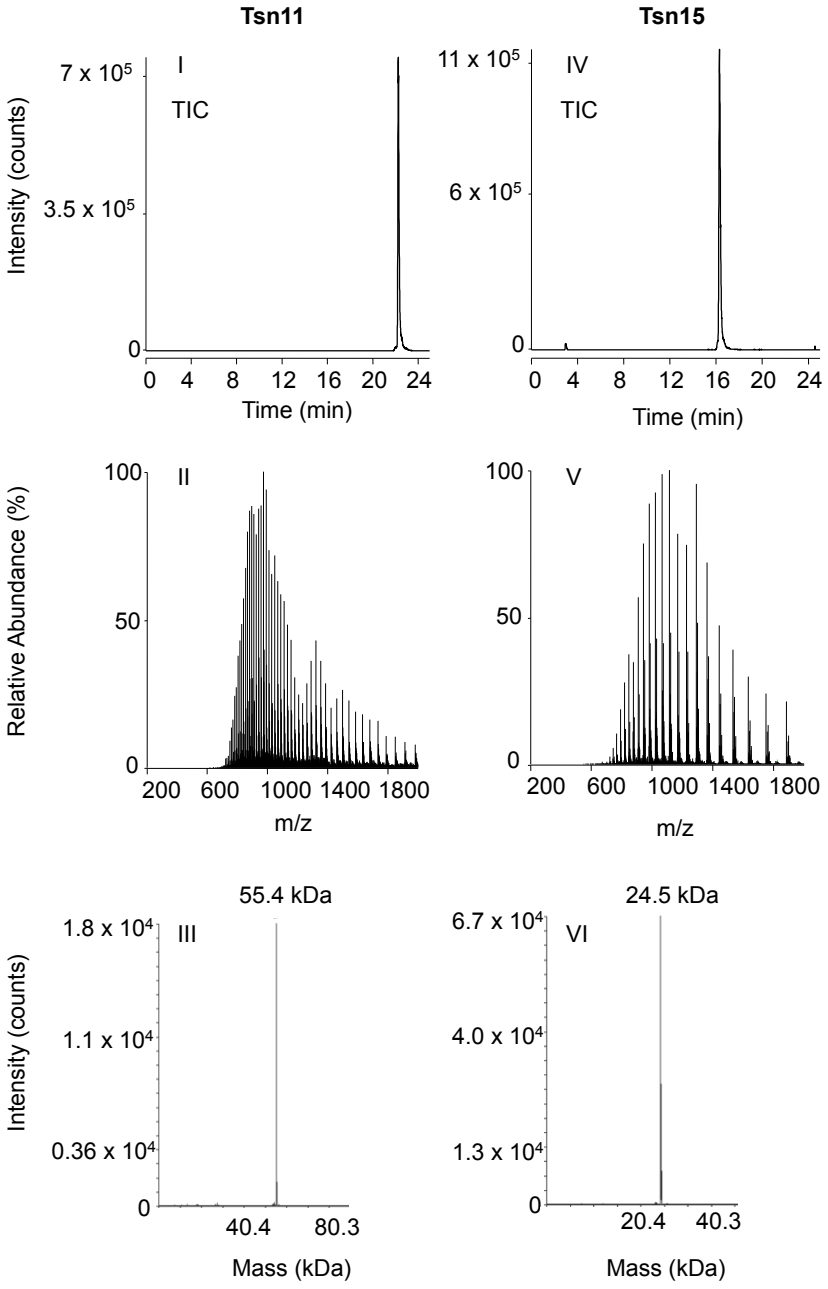
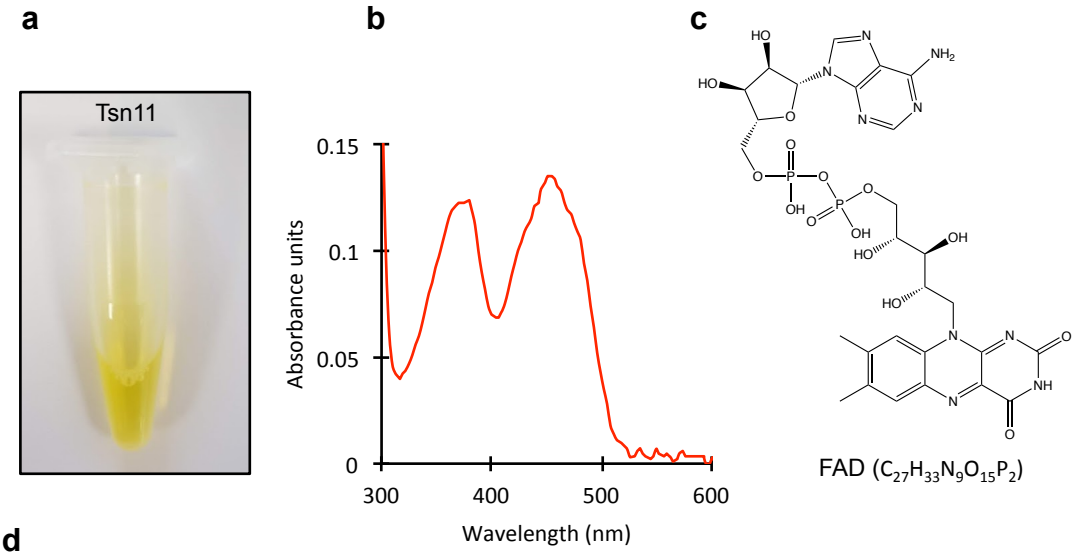


Figure 5.9 LC-MS of Tsn11 and Tsn15. (I) Total ion current (TIC) chromatogram of purified Tsn11. (II) Mass spectrum of Tsn11. (III) Deconvoluted mass spectrum of Tsn11. The detected z=1 mass of 55.4 kDa is identical to its predicted mass of 55.4 kDa. (IV) TIC chromatogram of purified Tsn15. (V) Mass spectrum of Tsn15. (VI) Deconvoluted mass spectrum of Tsn15. The detected z=1 mass of 24.5 kDa is consistent with the predicted mass of 24.4 kDa.

Tsn11 purified as a bright yellow protein with a UV_{MAX} of 450 nm, indicating it has a bound flavin prosthetic group (FAD or FMN) (**Figure 5.10**). To determine which prosthetic group was present, a small sample of Tsn11 was boiled to denature the protein and release the prosthetic group. HRMS was then used to identify the bound prosthetic group as FAD (as in PyrE3)[161] (**Figure 5.10d**). The stoichiometry in which FAD bound was investigated more in Chapter 6.



d

Name	Formula	Neutral Mass	Observed neutral mass	Observed m/z	Adduct	Error (ppm)
FAD	$C_{27}H_{33}N_9O_{15}P_2$	785.15713	785.1540	784.1467	−H	−4.3

Figure 5.10 The FAD prosthetic group of Tsn11. a, A sample of purified Tsn11 in a 1.5 mL microcentrifuge tube. **b,** UV/Vis absorbance spectrum of Tsn11. **c,** structure of chemical formula of FAD. **d,** Table of the predicted and detected HRMS of FAD collected from the supernatant of a boiled Tsn11 sample.

5.2.2 *in vitro* reconstitution of cyclohexane and tetrahydropyran ring formation in tetronasin biosynthesis

Purified Tsn11 and Tsn15 were used with T-16 in *in vitro* assays. The *in vitro* enzyme assays were performed at 30 °C for 1 h, after which the reactions were terminated and the products analysed by HPLC (**Figure 5.11**). When T-16 was incubated with Tsn11, the T-16 peak disappeared and a peak corresponding to a new metabolite labelled T-22, appeared. T-22 has the same apparent molecular weight as T-16 and tetronasin, but a unique

chromophore and retention time (discussed more in **Section 5.34**). In contrast, no conversion occurred when T-16 was incubated with Tsn15 alone. However, incubation of T-16 with both Tsn11 and Tsn15 converted T-16 into tetronasin. The identity of tetronasin was confirmed by the produced species having an identical retention time, HRMS, UV/Vis absorption spectrum, and MS^3 fragmentation pattern to a tetronasin standard (**Figure 5.12, Figure 5.13, Figure 5.14**). Incubating T-22 with Tsn15 also produced tetronasin, demonstrating that, as predicted by the analysis of the *S. longisporoflavus* ∆tsn11 ∆tsn15 mutant (**Figure 5.21**), Tsn11 precedes Tsn15. No conversion of T-16 was observed in the absence of enzymes. An alternative trace of this assay with better separation between T-22 and tetronasin (albeit with less defined peaks) is presented in **Appendix Figure 4**.

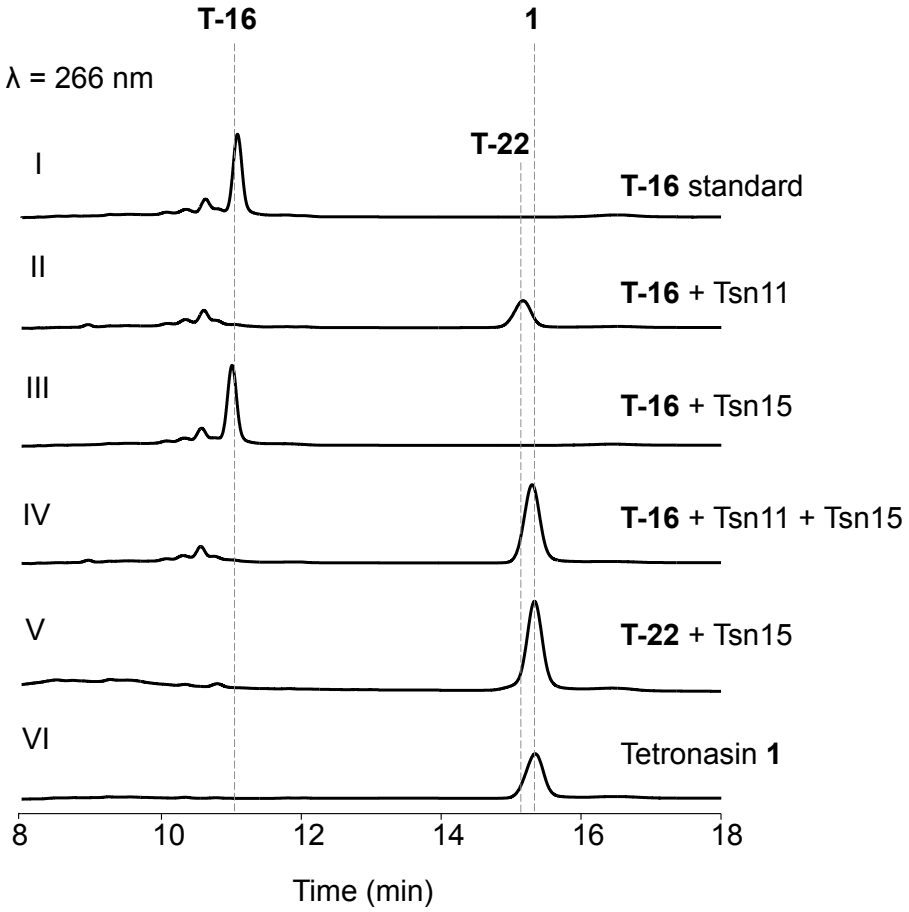


Figure 5.11 *In vitro* conversion of T-16 into tetronasin using purified Tsn11 and Tsn15. The *in vitro* conversion of T-16 into tetronasin **1** by Tsn11 and Tsn15 was monitored by HPLC analysis. (I) T-16 incubated in the absence of enzymes. (II) Conversion of T-16 into T-22 using Tsn11 (III) No reaction occurred when T-16 was incubated with Tsn15 alone. (IV) Conversion of T-16 into tetronasin **1** using Tsn11 and Tsn15. (V) Conversion of T-22 into tetronasin **1** using Tsn15. (VI) Tetronasin **1** standard. Data are representative of three independent experiments.

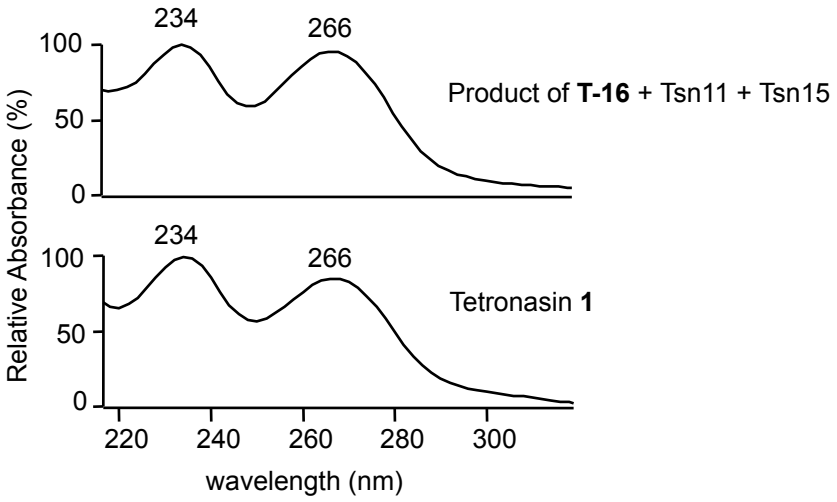


Figure 5.12 UV/Vis spectra comparison of the T-16 assay product and a tetronasin standard. (I) UV/vis spectra of the *in vitro* assay product of T-16 incubated with Tsn11 and Tsn15. UV_{max} = 234 nm, 266 nm (II) Chromophore of a tetronasin **1** standard UV_{max} = 234 nm, 266 nm.

Formula	Theoretical mass	Observed m/z	Adduct	Error (ppm)
$C_{35}H_{54}O_8Na$	625.3711	625.3691	Na^+	−3.10

Figure 5.13 HRMS data of the product produced by the T-16 + Tsn11 + Tsn11 assay. High-resolution mass spectrometry of the product of the T-16 + Tsn11 + Tsn15 assay confirmed that it has the same molecular formula as tetronasin.

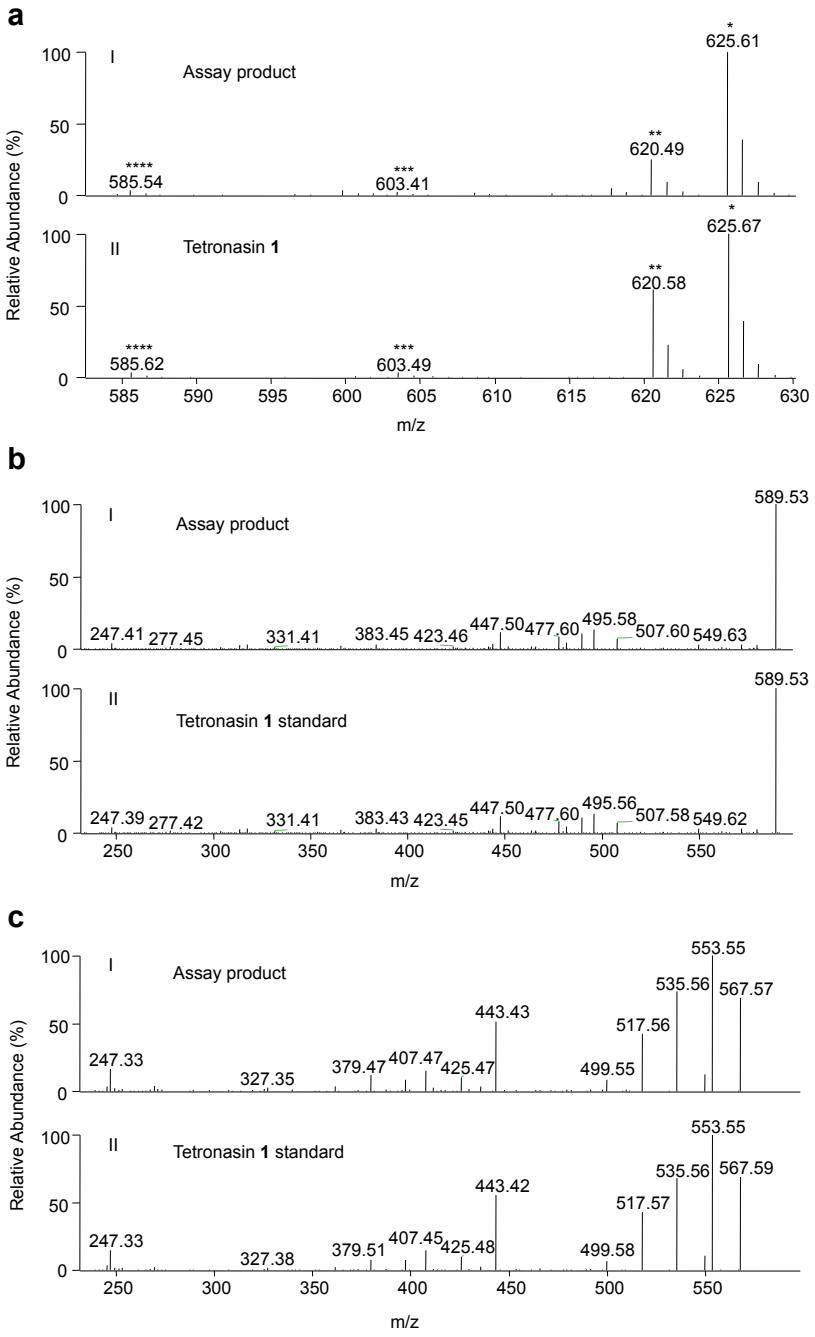


Figure 5.14 MS^3 comparison between the T-16 + Tsn11 + Tsn15 assay product and a tetronasin standard. a, (I) Mass spectrum of the product of T-16 + Tsn11 + Tsn15 assay; (II) a tetronasin **1** standard. *$[M+Na]^+$, **$[M+NH_4]^+$, ***$[M+H]^+$, ****$[M+H-H_2O]^+$. **b,** MS^3 mass spectra of the m/z = 625 $[M+Na]^+$ peak for (I) the product of the T-16 + Tsn11 + Tsn15 assay and (II) a tetronasin **1** standard. Normalised collision energy 35%. **c**, MS^3 mass spectra of the m/z = 603 $[M+H]^+$ peak for (I) the T-16 + Tsn11 + Tsn15 assay and (II) a tetronasin **1** standard.

Boiling Tsn11 abolished its activity, whereas Tsn15 remained active even after boiling for ten minutes (**Figure 5.15**). Adding EDTA to the assay buffer also did not inhibit the reaction, indicating chelatable metal ions were not required for the reaction to proceed.

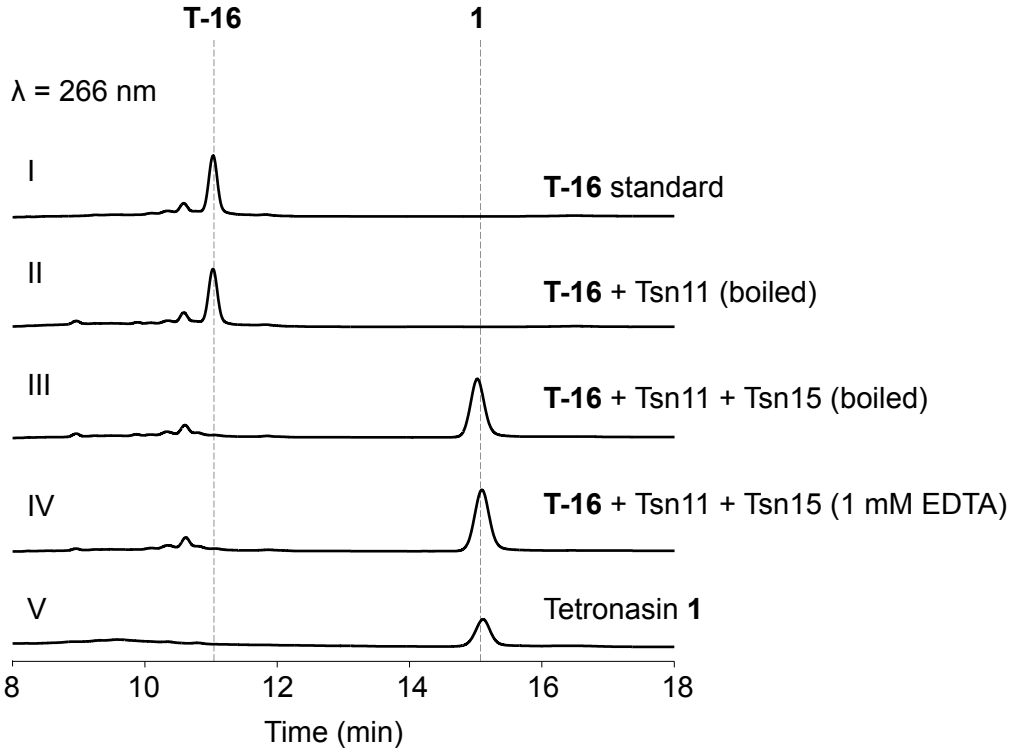


Figure 5.15 T-16 conversion assay in the presence of EDTA or boiled enzyme. (I) T-16 standard incubated in the absence of enzyme. (II) T-16 was not converted into T-22 when incubated with boiled Tsn11. (III) Conversion of T-16 into tetronasin **1** when incubated with Tsn11 and boiled Tsn15 (IV). The addition of 1 mM EDTA to the assay buffer did not interfere with the conversion of T-16 into tetronasin **1** by Tsn11 and Tsn15. (V) Tetronasin **1** standard. Data are representative of three independent experiments.

Taken together these data showed that Tsn11 and Tsn15 act in an enzymatic cascade to convert T-16 into tetronasin via the unknown intermediate T-22, closing both the cyclohexane and tetrahydropyran rings of tetronasin (**Figure 5.16**). The precise role of each enzyme in this process, i.e., if one enzyme forms both rings, or they each form one, was unknown. The structure of T-22 was the key piece of information required to answer this question.

Figure 5.16 Enzymatic cascade for the conversion of T-16 into tetronasin via T-22. T-16 is converted into tetronasin via two sequential enzyme-catalysed steps. Tsn11 converts T-16 into the unknown intermediate T-22, which is then converted into tetronasin by Tsn15.

5.2.3 Possible mechanisms of cyclohexane and tetrahydropyran ring closure

Prior to solving the structure of T-22 four main possible mechanisms were considered for cyclohexane and tetrahydropyran formation (**Figure 5.17**). Mechanism 1 was proposed by the author, Mechanisms 2 and 4 by Professor Gregory Challis (University of Warwick, UK), and Mechanism 3 by Dr. Finian Leeper (University of Cambridge, UK).

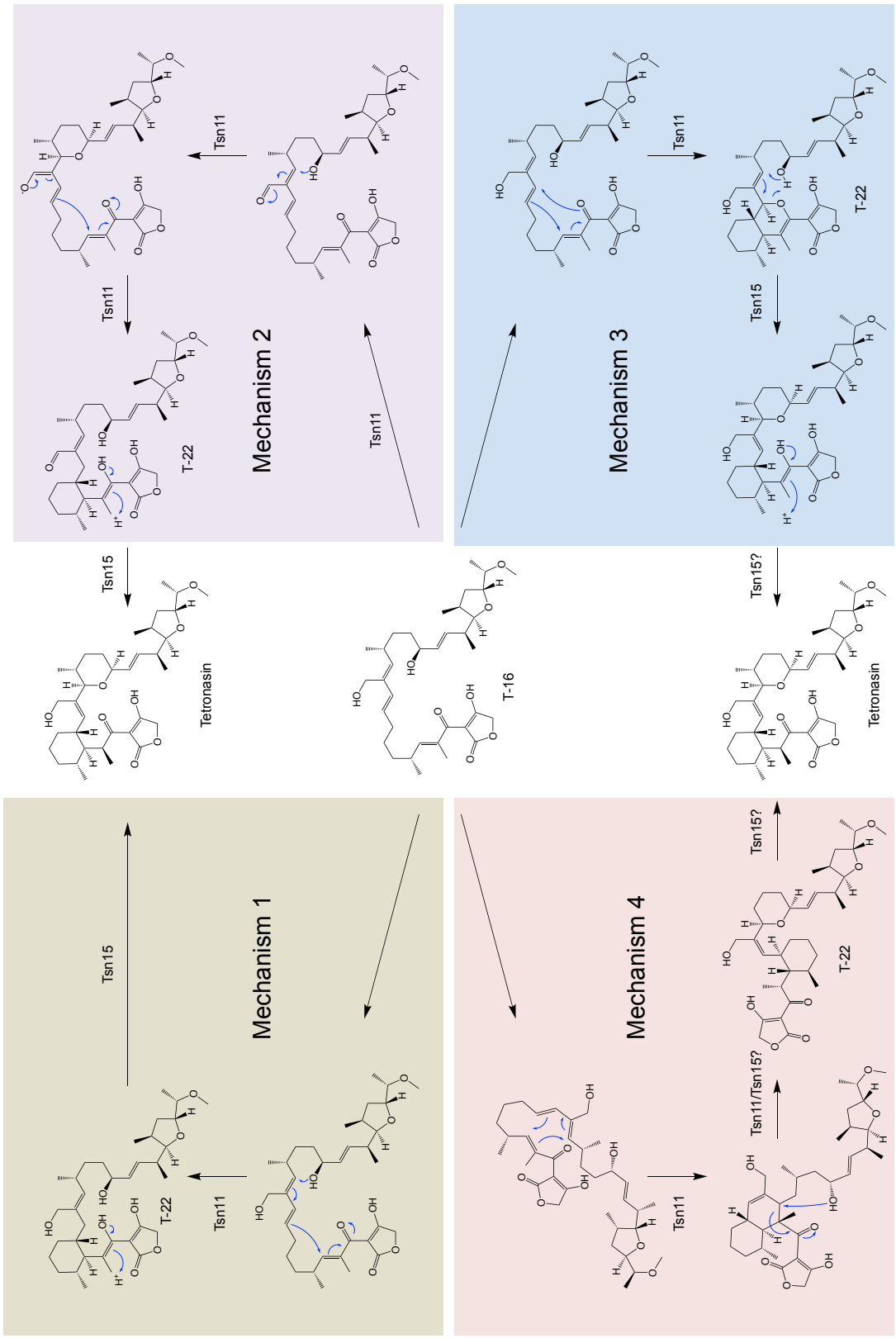


Figure 5.17 Mechanisms considered for conversion of T-16 into tetronasin by Tsn11 and Tsn15.

5.2.3.1 Mechanism 1

Mechanism 1 (**Figure 5.17**) resembles the previously proposed metal-catalysed cyclohexane and tetrahydropyran ring closure mechanism[175,177]. In this scheme Tsn11 catalyses cyclohexane and tetrahydropyran formation, leaving C3-C4 in the enol form. To form tetronasin Tsn15 catalyses a stereospecific keto-enol tautomerisation to form the C3 keto and establish the *S* configuration of C32. Such keto-enol tautomerisation can occur spontaneously, but the specific stereochemical outcome required to produce tetronasin may indicate that an enzyme is responsible. Such stereospecific keto-enol tautomerases have been previously characterised, such as oxaloacetate keto-enol tautomerase[275,276]. This mechanism gains plausibility when considering the equivalent enol intermediate for tetronomycin. Both tetronasin and tetronomycin would have very similar C3-C4 double bond enol intermediates, yet their C4 methyl groups end up having the opposite stereochemistry. A stereospecific keto-enol tautomerase enzyme could therefore be required for establishing this stereochemical difference (**Figure 5.18**). In the second total synthesis of tetronasin the cyclohexane and tetrahydropyran rings could be formed in a single step using an intermediate resembling T-16[180]. Although in that case the cyclohexane and tetrahydropyran ring stereochemistry was correct, the C4 stereochemistry was not, providing further evidence of enzymatic involvement[180].

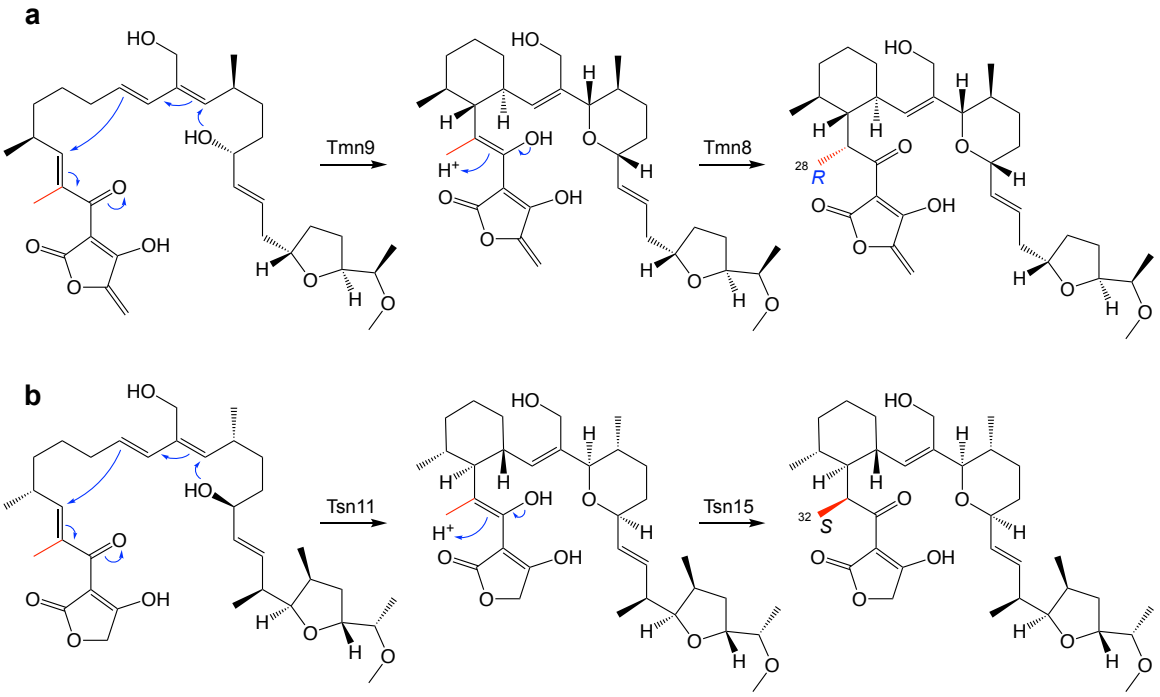


Figure 5.18 Possible mechanism for forming the R-configured C28 methyl group in tetronomycin. a, Under the Mechanism 1 proposal for tetrahydropyran and cyclohexane formation, Tmn8 is responsible for catalysing the stereospecific keto-enol tautomerisation to form the *R*-configured C28 methyl group. **b,** Proposed formation of the *S*-configured C32 methyl group in tetronasin under Mechanism 1.

5.2.3.2 Mechanism 2

Mechanism 2 (**Figure 5.17**) proposes that Tsn11 is able to bind NADPH and has FAD-dependent monooxygenase activity. In this scheme Tsn11 oxidises the C30 hydroxyl to an aldehyde. The aldehyde activates the C13 for attack by the C17 hydroxyl, closing the tetrahydropyran ring. Next, while still bound to Tsn11, the cyclohexane ring forms. Tsn11 then reduces the aldehyde back to the C30 hydroxyl. A similar oxidation-reduction cycle is found in enzymes such as dehydroquinate synthase[277]. Tsn15 then catalyses the same stereospecific keto-enol tautomerisation proposed in Mechanism 1 to form the *S* configured C32 methyl group. One issue this mechanism faces is that formation of an aldehyde at C30 was previously ruled out[176]. In that study, two deuterium atoms were incorporated at C30 by feeding *S. longisporoflavus* deuterium-labelled propionate[176]. Retention of deuterium at this site is not expected if the hydroxyl was oxidised to an aldehyde then reduced back to an alcohol, at which point hydrogen atoms from the solvent would be incorporated [176]. However, this mechanism could be occurring if the same proton removed by FAD in the oxidation reaction were protected from solvent and returned in the reduction reaction. The

FAD prosthetic group of Tsn11 would be essential for Tsn11 activity if this mechanism were correct.

5.2.3.3 Mechanism 3

In Mechanism 3 (**Figure 5.17**) Tsn11 catalyses an inverse-electron-demand hetero-Diels-Alder/[4+2] cycloaddition ($_{INV}$HDA) reaction (**Figure 5.19**) to form an oxadecalin intermediate, closing the cyclohexane ring. Tsn15 then catalyses tetrahydropyran ring formation and fragmentation of the oxadecalin to form tetronasin. This mechanism has the

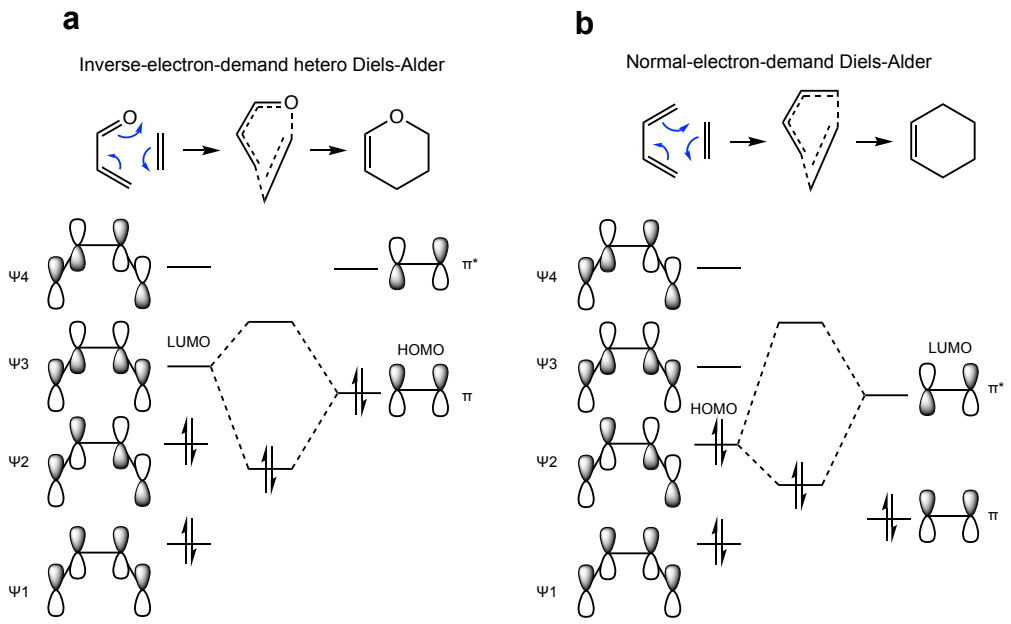


Figure 5.19 Inverse-electron-demand hetero-Diels-Alder reaction. a, In an inverse-electron-demand hetero-Diels-Alder reaction an electron withdrawing group on the diene results in the LUMO of the diene and HOMO of the dieneophile (alkene) being closer in energy than the HOMO of the diene and LUMO of the dieneophile. b, Energy level diagram for a normal-electon-demand Diels-Alder reaction.

advantage that both Tsn11 and Tsn15 would be catalysing pericyclic reactions like their homologues in spirotetronate/tetramate biosynthesis[157,158,161] (and AOC2) in the case of Tsn15[278]). Other examples of enzymes catalysing apparent $_{INV}$HDA reactions have also been reported[279,280], so there is some precedent. In this mechanism the keto-enol tautomerisation at C3-C4 is either spontaneous or also catalysed by Tsn15.

5.2.3.4 Mechanism 4

Mechanism 4 (**Figure 5.17**) is *prima facie* the least likely mechanism. In this scheme Tsn11 catalyses carbocyclic decalin formation, just like its homologue PyrE3 from pyrroindomycin biosynthesis[161]. Tsn11 or Tsn15 then catalyses tetrahydropyran formation, requiring attack of the C18 hydroxyl onto a sp^3 hybridised carbon, and fragmentation of the decalin ring. The resulting molecule resembles tetronasin except the C11-C12 double bond is in the *cis* configuration rather than the *trans.* To form tetronasin from this structure a *cis-trans* isomerisation is required, possibly catalysed by Tsn15.

5.2.4 Purification and structural characterisation of T-22

To gain insight into the mechanism by which Tsn11 and Tsn15 convert T-16 into tetronasin, it was necessary to purify and structurally characterise the intermediate. T-22 has a distinctive chromophore (λ = 254 nm, 288 nm, MeOH) compared to T-16 and tetronasin. The mass spectrum of T-22 resembles tetronasin more than T-16, the major ion no longer being $[M+H-H_2O]^+$ (m/z = 585) (**Figure 5.20**). The MS^3 fragmentation pattern of T-22 was also very similar to that of tetronasin (**Appendix figure 5**).

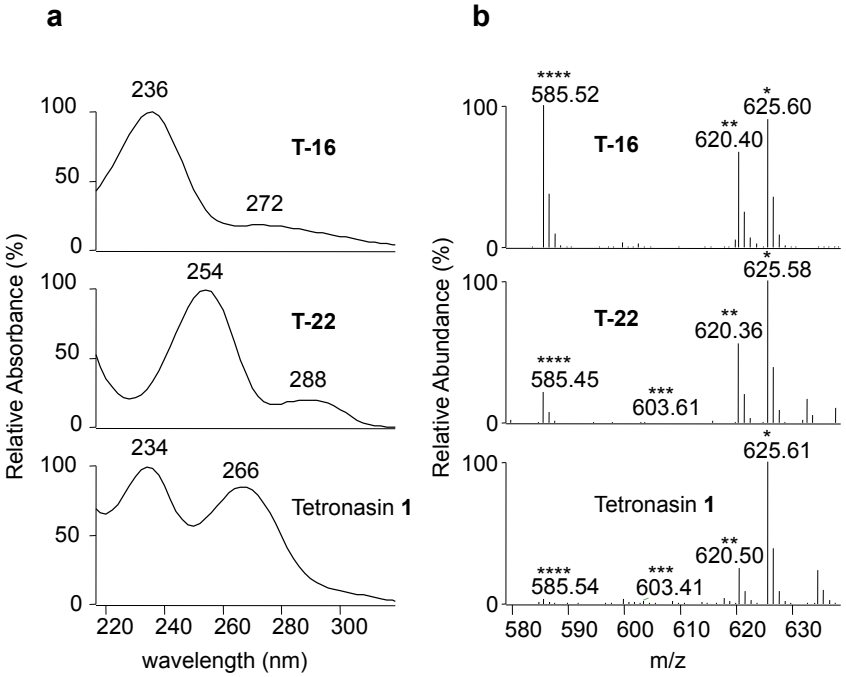


Figure 5.20 UV/Vis spectrum and mass spectrum of T-22. (**A**) The UV/vis absorption specta of T-16 (λ_{MAX} = 236 nm, 272 nm), T-22 (λ_{MAX} = 254 nm, 288 nm), and tetronasin (λ_{MAX} = 234 nm, 266 nm). (**B**) Mass spectrum of T-16, T-22, and tetronasin. *$[M+Na]^+$, **$[M+NH_4]^+$, ***$[M+H]^+$, ****$[M+H-H_2O]^+$.

T-22 was clearly related to T-21, the metabolite that accumulated in the *S. longisporoflavus* ∆tsn15 mutant (**Figure 5.21**). Like T-21, the major detected ions of T-22 were m/z = 538 and m/z = 543. However, the putative $[M+H]^+$, $[M+NH_4]^+$, and $[M+Na]^+$ of T-22 were not shifted two mass units greater as observed for T-21.

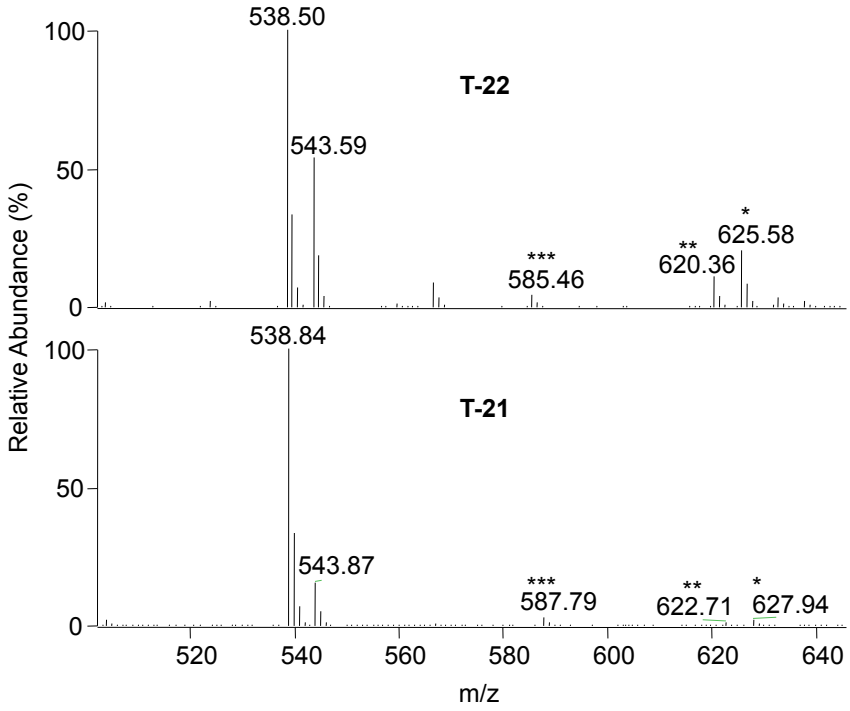


Figure 5.21 Comparison of T-22 and T-21 mass spectra. Top: mass spectrum of T-22, produced from an *in vitro* assay containing T-16 and Tsn11. Bottom: mass spectrum of T-21, produced by the *S. longisporoflavus* Δtsn15 mutant. *$[M+Na]^+$, **$[M+H]^+$, ***$[M+NH_4]^+$.

Given their apparent relatedness, it was tested if T-21, like T-22, is also a substrate of Tsn15. Rather than try to purify T-21 from the crude organic extract of *S. longisporoflavus* Δtsn15, the extract itself was incubated with Tsn15 and analysed by HPLC. However, no tetronasin production was detected from this assay (**Figure 5.22**). To rule out other metabolites in the crude extract inhibiting Tsn15, the *S. longisporoflavus* Δtsn15 crude extract was supplemented with T-16. Adding Tsn11 and Tsn15 to this supplemented extract produced tetronasin, demonstrating that no component of the crude extract inhibited Tsn11 or Tsn15. T-21 can therefore not be converted into tetronasin by Tsn15. The exact identity of T-21 remains unknown, but it is potentially a breakdown product of T-22 formed during the fermentation of *S. longisporoflavus* Δtsn15.

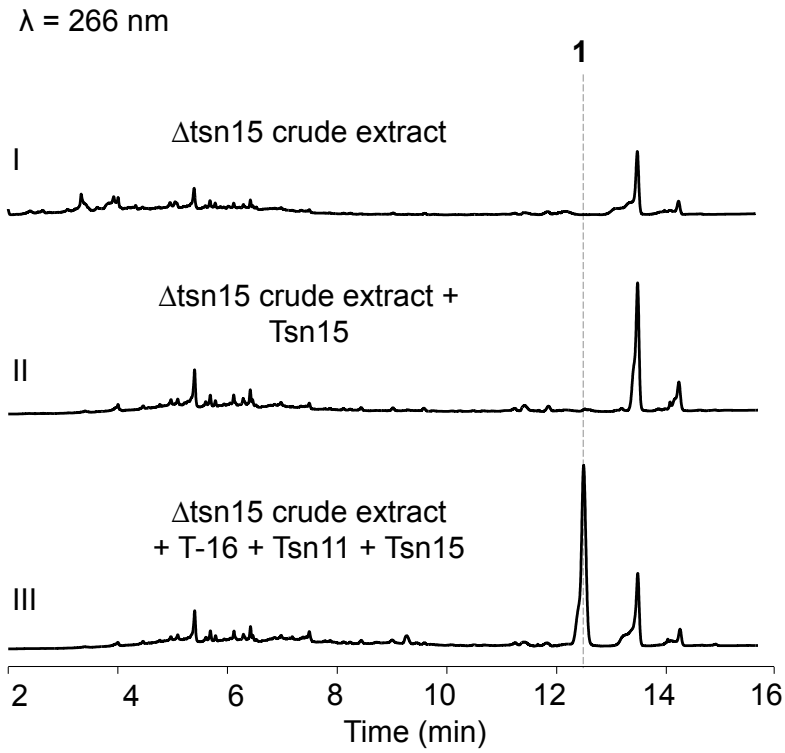


Figure 5.22 Tsn15 assay using crude organic extract from the *S. longisporoflavus* Δtsn15 mutant as a substrate. (I) Chromatogram of the crude organic extract from the *S. longisporoflavus* Δtsn15 mutant. (II) No tetronasin production was observed when Tsn15 was incubated with the *S. longisporoflavus* Δtsn15 crude organic extract. (III) Production of tetronasin by supplementing *S. longisporoflavus* Δtsn15 crude organic extract with T-16, Tsn11, and Tsn15.

To produce sufficient T-22 for NMR analysis, the conversion of T-16 by Tsn11 was scaled up, after first optimising the pH of the assay to maximise production of T-22. To achieve this, T-16 was incubated with Tsn11 in buffers ranging in pH from 6.0-9.0, and T-22 formation was monitored by HPLC (**Figure 5.23**). Tsn11 was found to be most active in the pH range 7.0-8.0, and accordingly the large-scale Tsn11 assay was conducted at pH 7.9.

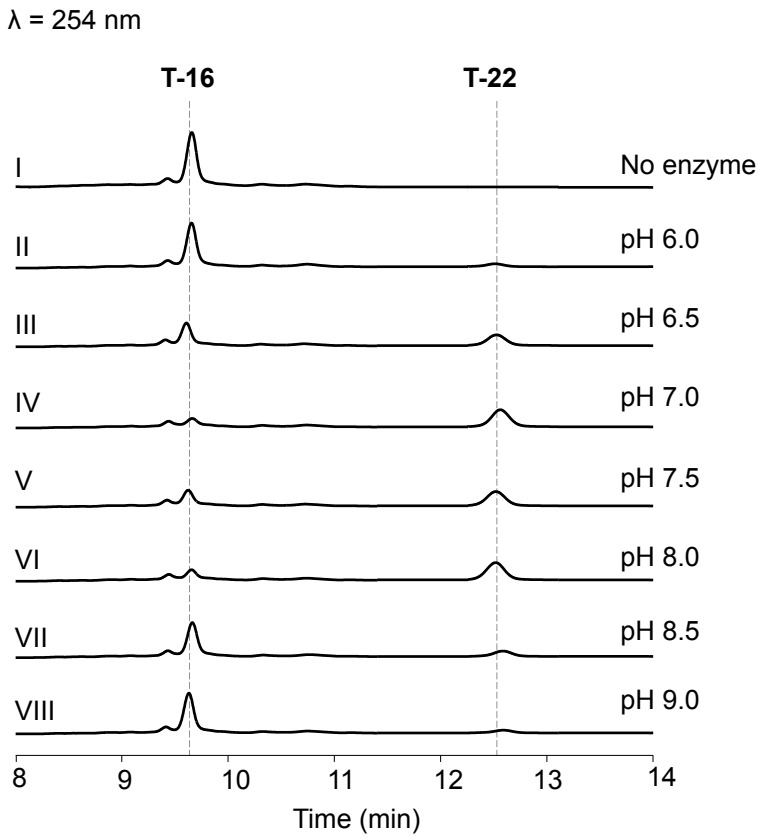


Figure 5.23 pH optimisation of the Tsn11 assay. T-16 and Tsn11 were incubated together in a range different pH buffers. The assay was terminated at 10 minutes, before complete conversion of T-16 into T-22 had occurred in any of the conditions. (I) T-16 no enzyme control. (II) T-16 + Tsn11 assay run in MES pH 6.0 buffer. (III) T-16 + Tsn11 assay run in pH MES pH 6.5 buffer. (IV) T-16 + Tsn11 assay run in TrisCl pH 7.0 buffer. (V) T-16 + Tsn11 assay run in TrisCl pH 7.5 buffer. (VI) T-16 + Tsn11 assay run in TrisCl pH 8.0 buffer. (VII) T-16 + Tsn11 assay run in TrisCl pH 8.5 buffer. (VIII) T-16 + Tsn11 assay run in TrisCl pH 9.0 buffer.

The initial attempt at scaling up the Tsn11 reaction, using the T-16 purified from the 100 initial plates, was unsuccessful. The amount of T-22 isolated (roughly 1 mg) was not sufficient to obtain clear NMR spectra. It was therefore necessary to increase the amount of T-16 used in the assay. *S. longisporoflavus* Δtsn11 was grown on 300 more tsn-medium-B plates. From this new batch, about 25 mg of T-16 was isolated as before, 15 mg of which was incubated with Tsn11 for conversion into T-22. The large-scale incubation was conducted at 30 °C for 1 h, after which T-22 was extracted using ethyl acetate and purified using semi-preparative HPLC (**Figure 5.24**).

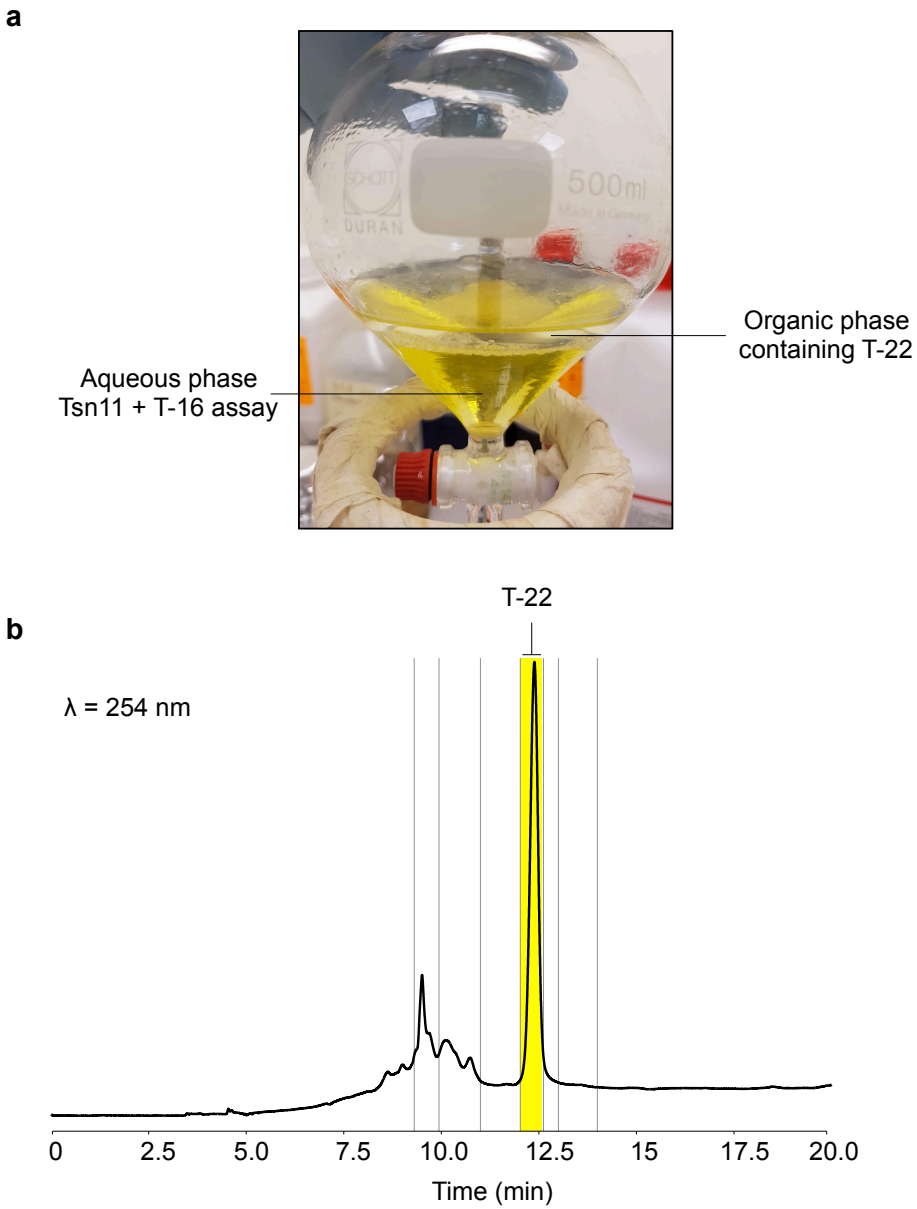


Figure 5.24 Production and purification of T-22 for NMR. a, Tsn11 was incubated with T-16 for 1 h to achieve complete conversion to T-22. T-22 was then extracted from the assay using ethyl acetate. When mixed with water ethyl acetate formed an upper organic phase **b,** Representative example of T-22 purification using a semi-preparative HPLC fraction collector. The grey lines show the boundaries of the different fractions automatically collected by the instrument. The fourth fraction (highlighted in yellow) collected T-22. The other peaks represent unrelated species.

The second attempt at the scaled-up Tsn11 assay was more successful than the first, producing ca. 3.5 mg of purified T-22. 1D and 2D NMR experiments including 1H, 13C, COSY, HSQC, HMBC, TOCSY, and NOESY experiments were then performed (**Appendix Figure 16-21, Appendix Figure 2**). NMR data was again analysed in collaboration with Dr. Finian Leeper (University of Cambridge). The NMR data enabled the structure of T-22 to be solved, revealing that it still lacked the tetrahydropyran ring, but contained a closed

cyclohexane ring (the C5-C10 bond had formed) as part of an unexpected oxadecalin moiety (**Figure 5.25**).

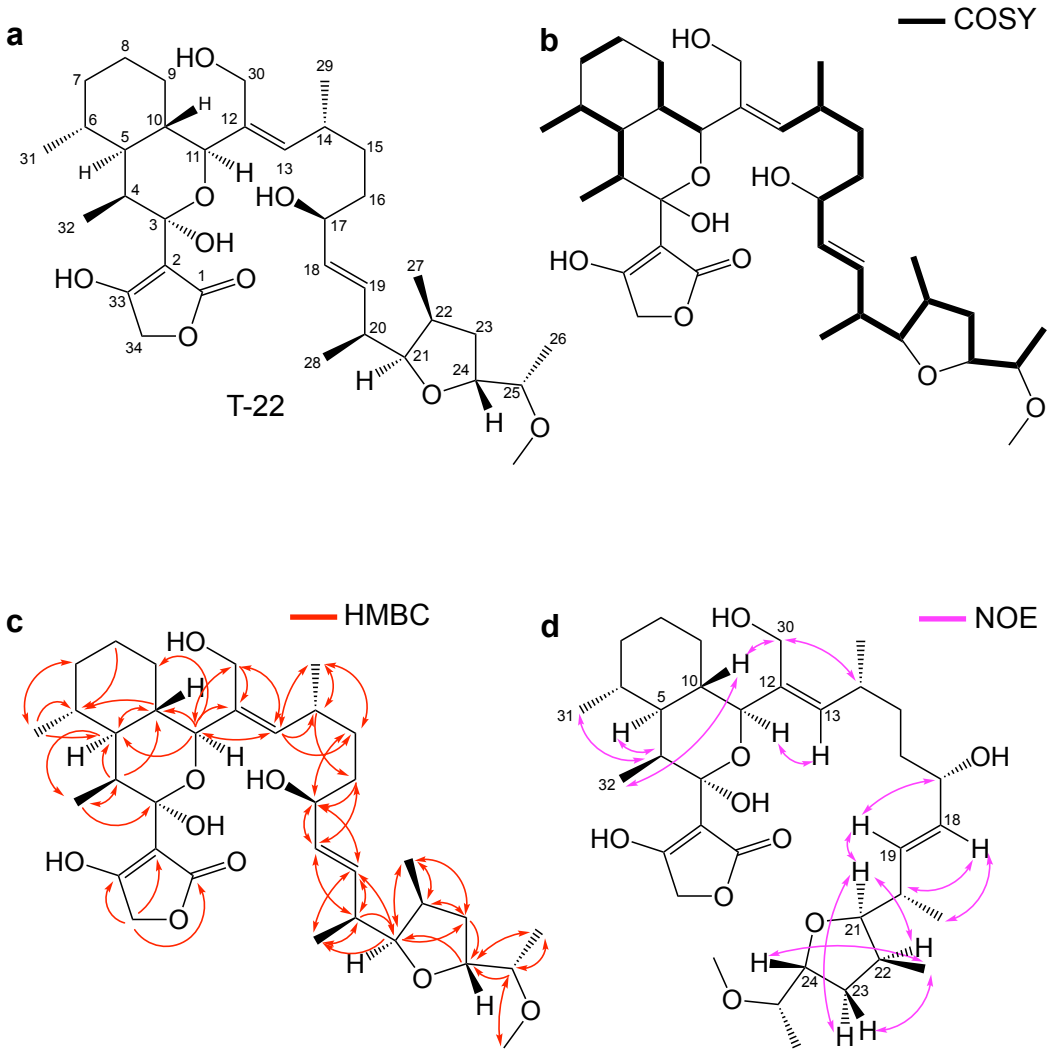


Figure 5.25 The structure of T-22 as determined by NMR spectroscopy. a, The structure of T-22 with atom labels. **b,** ^{1}H-^{1}H COSY correlations for T-22. **c,** ^{1}H-^{13}C HMBC correlations of T-22. Arrows point from proton to carbon. **d,** Selected ^{1}H-^{1}H NOESY correlations of T-22.

As in T-16, the signals from C17 to C28, including the methoxy group, were very similar to those in tetronasin[174,181], so could be readily assigned from the COSY and HMBC connections. The configuration of the double bonds could be determined as *trans* by the NOESY, as could the relative stereochemistry of the tetrahydrofuran ring (the same as T-16 and tetronasin). Unlike T-16, however, which had six protons in the olefinic regions of the ^{1}H spectrum, T-22 only had three. The three fewer olefinic protons immediately

suggested that the cyclohexane ring had closed. Two of the olefinic peaks correspond to the protons attached to C18 and C19 in the tetrahydrofuran portion of the molecule. COSY and HMBC correlations could determine that the third olefinic peak is the C13 proton (showing an HMBC correlation to the C14 methyl group). The TOCSY NMR spectrum further supported this assignment, showing all three olefinic protons (C13, C18, and C19) are in the same spin system as each other. The ^{1}H chemical shift of C11 proton (3.81 ppm) shows that C11 is sp^{3} hybridised and has an oxygen atom attached. This oxygen appears to be cyclised onto C3 to form a cyclic hemiacetal oxadecalin. The evidence for this is the chemical shift of C-3 (104.5 ppm). If the C11 hydroxyl were not cyclised onto C3, then the major component would contain a keto group at C3 and its ^{13}C chemical shift would be further downfield at ca. 190-200 ppm[174,274] (**Figure 5.26**).

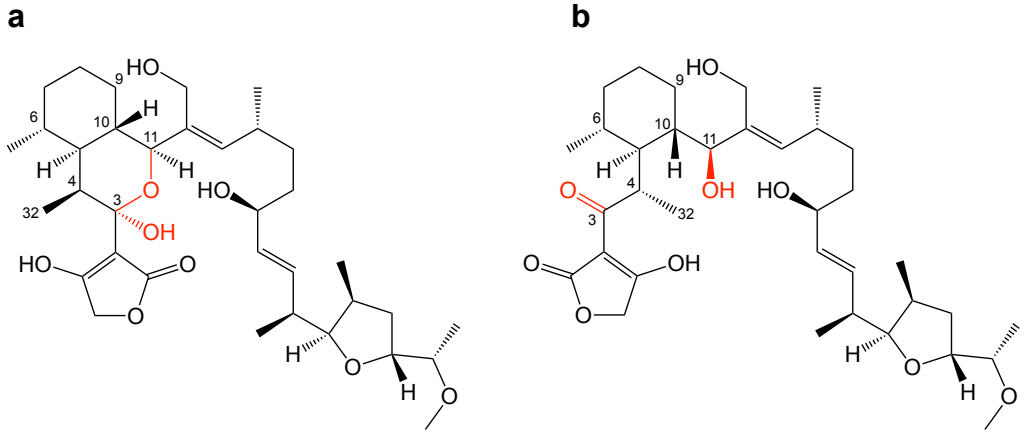


Figure 5.26 Closed and open forms of T-22. a, the proposed major component structure of T-22 where a cyclic hemiacetal has formed by the C11 hydroxyl cyclising onto C3. The measured ^{13}C shift for C3 is 104.5 ppm. **b,** The alternative structure of T-22 where cyclic hemiacetal formation has not occurred, resulting in a C3 keto group. The C3 keto structure is not supported as the ^{13}C chemical shift would be expected at around 190-200 ppm[174,274], which is not observed.

Furthermore, the ^{13}C shift of C33 is 179.99 ppm, whereas if C3 were a ketone it would be expected at ca. 190 ppm[174,274]. The final piece of evidence for the cyclic hemiacetal structure came from the NOESY spectrum. The NOESY revealed that T-22 was in equilibrium with several minor components, resulting in negative NOE peaks (standard NOEs show as positive peaks) due to saturation transfer between equivalent protons in these major and minor components (**Figure 5.27a**). The largest saturation transfer difference was for the H4 (δ 3.81 major to δ 4.67 minor) and H11 protons (δ 2.355 major to δ 4.05 minor), strongly suggesting the reaction occurring is ring opening of the hemiacetal

to form a C3 keto and C11 hydroxyl (**Figure 5.27b**). Smaller saturation transfer signals were also present for protons adjacent to H4 and H11. The saturation transfer peaks of the two C34 protons (δ 4.62 in the major hemiacetal component), indicate that there are actually two minor C3 keto components, likely corresponding to the two different rotomers of the C2-C3 bond (singlets at δ 4.46 and δ 4.42). By integrating these peaks the ratio of the different components was estimated as 75% cyclic hemiacetal (H-34 δ 4.62), 15% ketone rotamer 1 (H-34 δ 4.46), and 10% ketone rotamer 2 (H-34 δ 4.42).

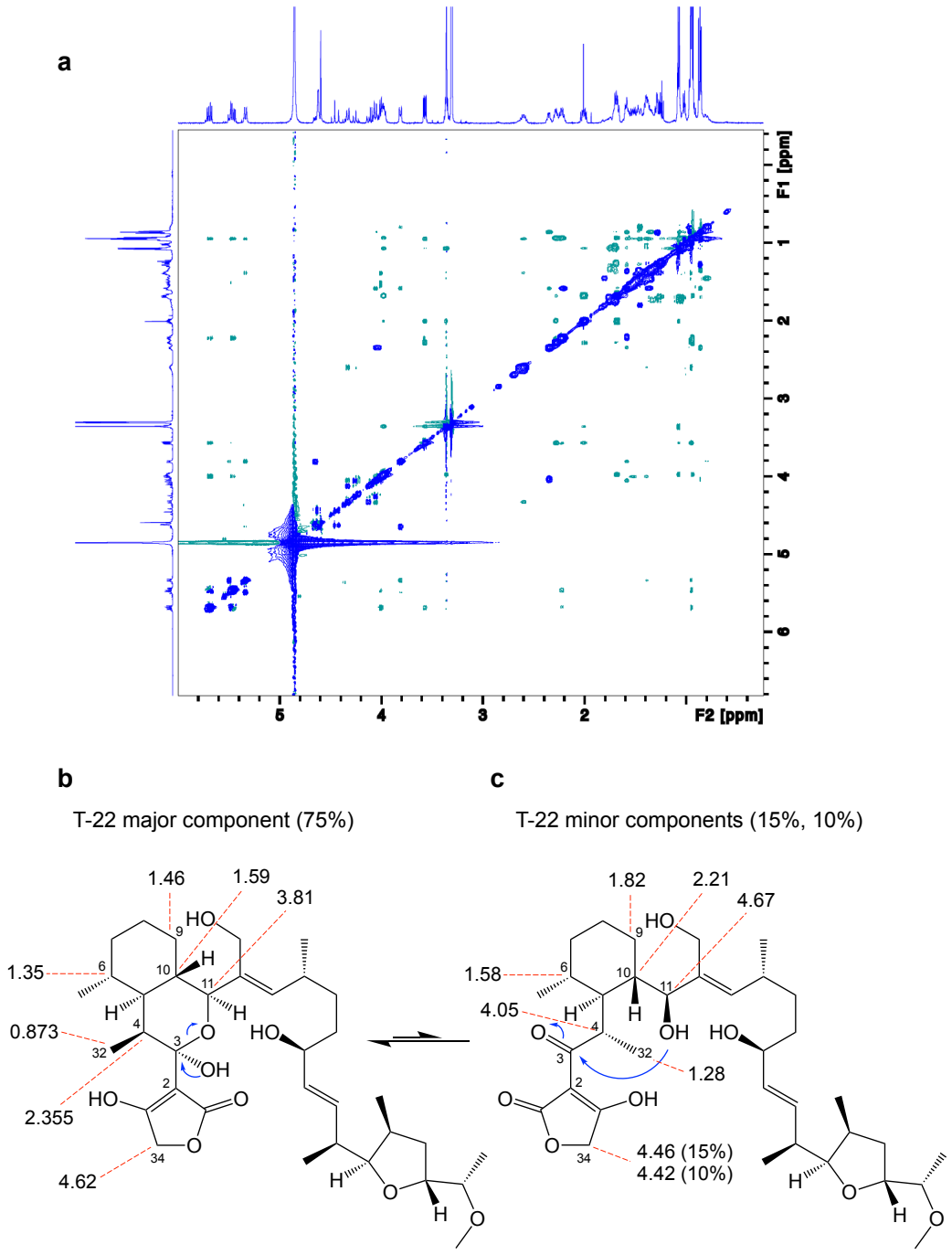


Figure 5.27 Saturation transfer signals in T-22. a, NOESY spectrum of T-22. Standard positive NOEs are presented as teal peaks. Negative peaks representing saturation transfer occurring between equivalent protons in the major and minor species are presented in dark blue. **b,** The major component (75%) of T-22 as a cyclic hemiacetal oxadecalin. The chemical shifts of protons in equilibrium with minor component forms are listed. **c,** The minor components of T-22 was determined by saturation transfer signals in the NOESY spectrum. The largest differences in ^{1}H chemical shifts are between the protons attached to C4 and C11. T-22 has two minor components due to rotation around the C2-C3 bond giving two rotomers. The ratio of these different components can be estimated as 75% cyclic hemiacetal, 15 % ketone rotamer 1 (H-34 δ 4.46), and 10% ketone rotamer 2 (H-34 δ 4.42).

The stereochemistry of H-11 was determined by its coupling constant to H-10 (*J* 9.7 Hz) indicating the two protons are *trans* axial-axial to each other. As H-10 is part of the cyclohexane ring its stereochemistry is highly unlikely to differ from that found in tetronasin. H-10 also has an NOE to the C32 methyl group, the stereochemistry of which is also almost certainly the same as in tetronasin. Taking these data together H-11 can be confidently assigned as being in an axial position on the oxadecalin ring and *trans* to H-10. The axial configuration of H-11 is also sterically favourable, as it places the bulky C12-C30 part of the compound in an equatorial position.

The stereochemistry of the C3 hydroxyl group could not be determined from the NMR, but there are good reasons for concluding that it is in the axial form as depicted. As the cyclic hemiacetal is reversibly opening and closing, the two stereoisomers of the ring-closed form should be in equilibrium. The isomer where the C3-OH is in the axial position should be preferred by the anomeric effect. The bulky tetronate is also sterically preferred in the equatorial position. As the second isomer is not observed it must be present at too low a percentage in the equilibrium (**Figure 5.28**).

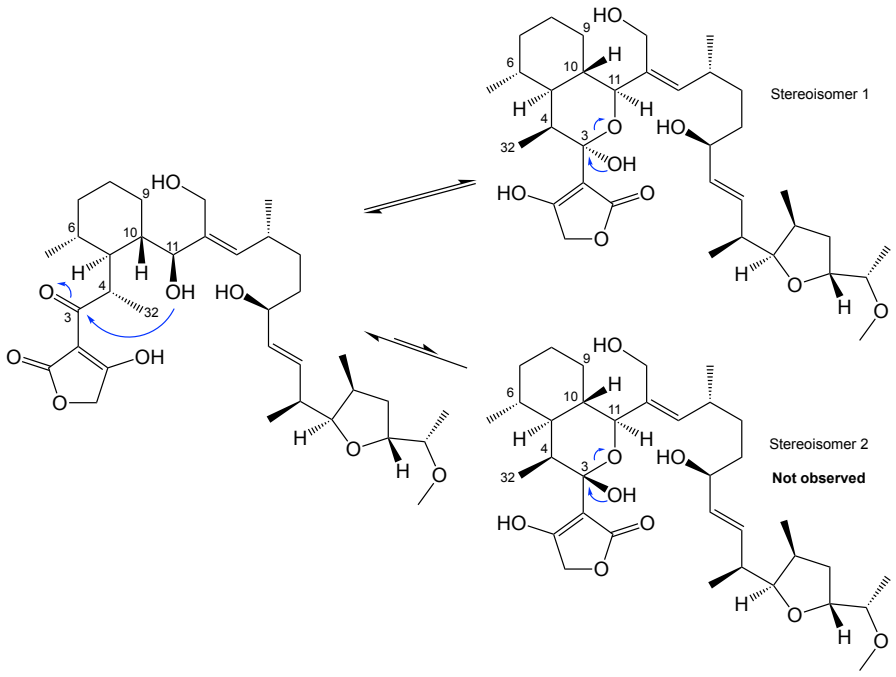


Figure 5.28 Stereoisomers of the T-22. The cyclic hemiacetal of T-22 had two possible stereoisomers. Only stereoisomer 1 was observed, likely due to the C3-hydroxyl being in the axial position, favourable due to the anomeric effect and the bulky tetronate moiety being in an equatorial position.

The additional hydroxyl group should make T-22 18 mass units heavier than T-16 or tetronasin. However, HRMS analysis could not detect this additional hydroxyl, its mass corresponding to the same molecular formula as the tetronasin sodium adduct ($C_{35}H_{54}O_8Na$) (**Figure 5.29**). However, hemiacetals are known to be unstable in mass spectrometers and readily fragment. It is therefore proposed that the detected mass corresponds to the $[M+Na-H_2O]^+$ ion. Having solved the structure of T-22, the enzymatic cascade converting T-16 into tetronasin, forming four stereocenters in the process, was fully resolved (**Figure 5.30**). Tsn11 converts T-16 into the oxadecalin T-22, establishing the final stereochemistry of C4, C5, and C10. Three chemical transformations are required to convert T-22 into tetronasin: dehydration of the cyclic hemiacetal, tetrahydropyran formation, and breakage of the C11-O bond (not necessarily in that order). Tsn15 appears to catalyse all three of these transformations and establish the stereochemistry of C13 in the process. Possible reaction mechanisms of Tsn11 and Tsn15 are discussed in the next section.

Formula	Theoretical mass	Observed m/z	Adduct	Error (ppm)
$C_{35}H_{54}O_8Na$	625.3711	625.3698	Na^+	−2.02

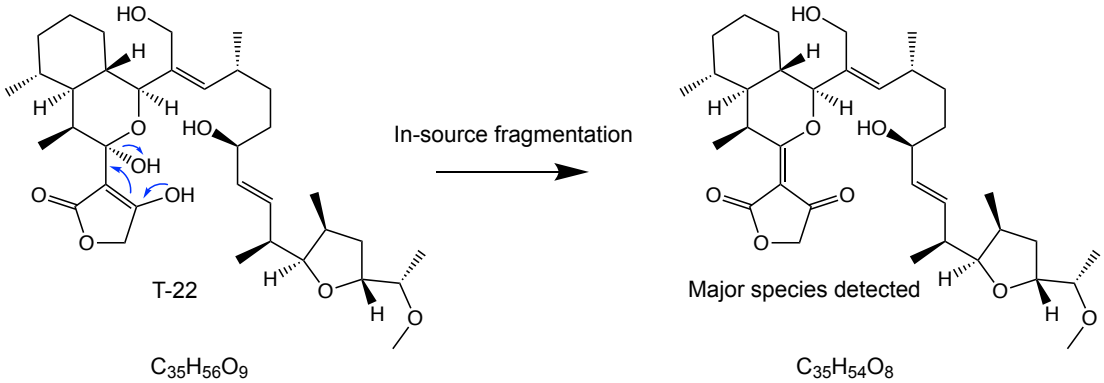


Figure 5.29 HRMS of T-22. The HRMS of T-22 indicated that its molecular weight was the same as tetronasin and T-16 ($C_{35}H_{54}O_8Na$). The additional hydroxyl group of the hemiacetal was not detected by the mass spectrometer. As hemiacetals are known to readily fragment in mass spectrometers, it is therefore likely that that measured mass represents the $[M+Na-H_2O]^+$ ion.

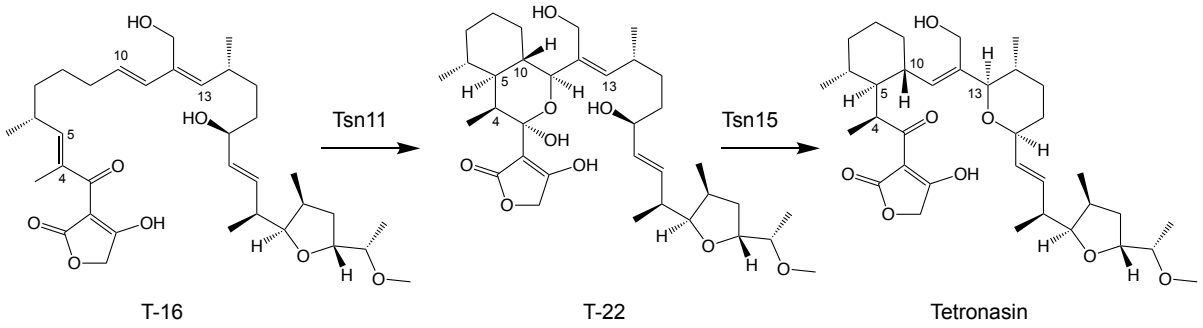
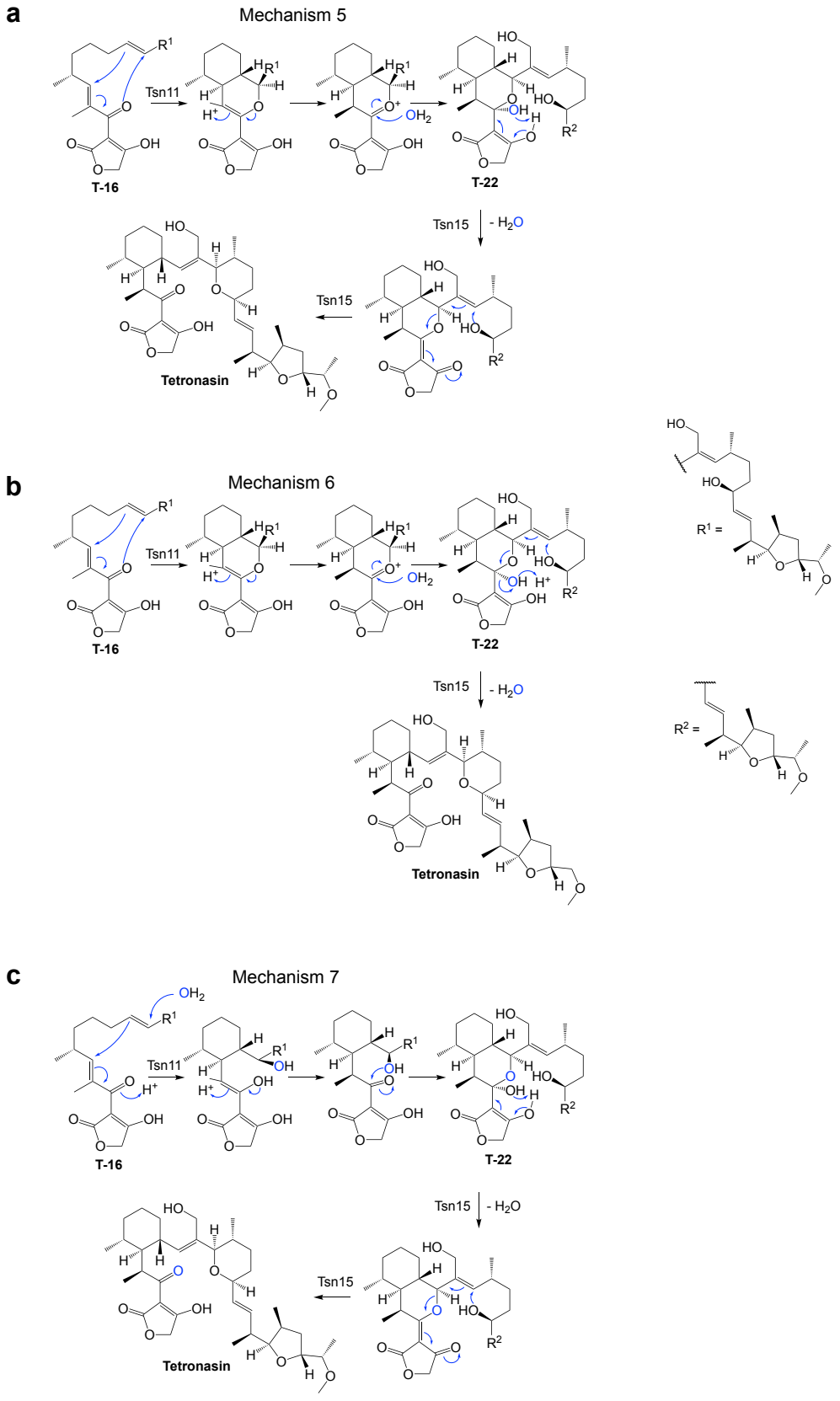


Figure 5.30 Formation of the cyclohexane and tetrahydropyran rings of tetronasin. The precyclisation substrate T-16 is converted into T-22 by Tsn11. The cyclohexane ring is closed in T-22 as part of a larger oxadecalin ring system. T-22 is then converted to tetronasin by Tsn15. Tsn15 catalyses formation of the tetrahydropyran ring, dehydration of the hemiacetal hydroxyl, and fragmentation of the oxadecalin. Together, Tsn11 and Tsn15 appear to establish the stereochemistry at C4, C5, C10, and C13 in tetronasin.

5.2.5 Possible mechanisms of enzyme-catalysed cyclohexane and tetrahydropyran formation

Given the structure of T-22, there are at least six alternative mechanisms that could account for the conversion of T-16 into tetronasin via T-22 by Tsn11 and Tsn15 (**Figure 5.31**). Some mechanisms are more likely than others, as shall be discussed. Tsn11 could convert T-16 into T-22 via an $_{INV}$HDA reaction (Mechanisms 5, 6, 9, and 10), as first proposed in Mechanism 3 (**Figure 5.17**). The resulting dihydropyran product must then tautomerise to the keto form, followed by attack by water onto C3 to form T-22. Alternatively, Tsn11 could catalyse attack of water onto C11, closing the cyclohexane ring in a stepwise reaction (Mechanisms 7 and 8).


a
Mechanism 5
Tsn11
T-16
T-22
Tsn15
- H2O
Tetronasin
b
Mechanism 6
R1 =
R2 =
c
Mechanism 7

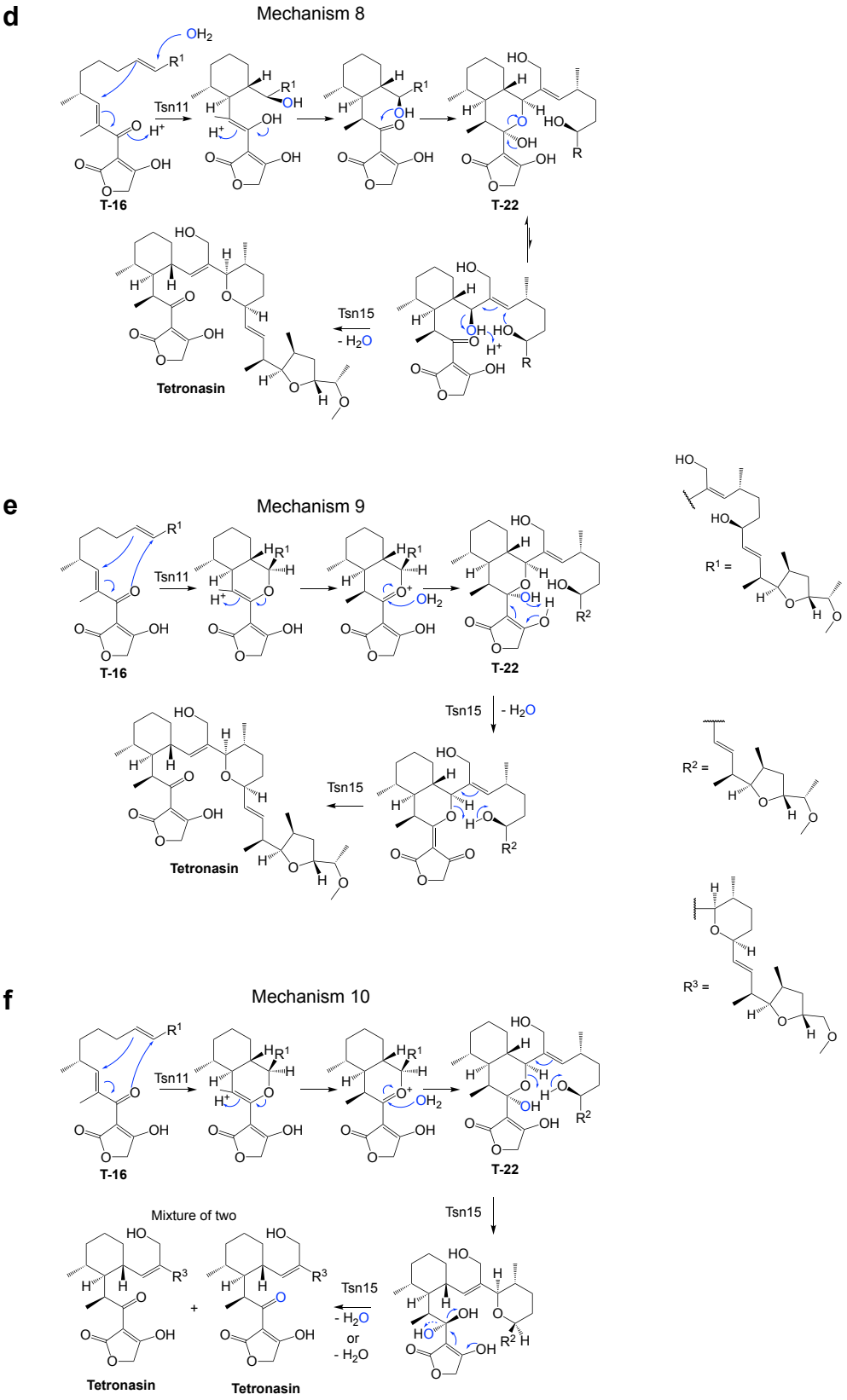


Figure 5.31 Possible mechanisms for the conversion of T-16 to tetronasin via T-22. At least six different mechanisms (Mechanism 5-10) are possible for the conversion of T-16 to tetronasin via T-22 by Tsn11 and Tsn15. Oxygen atoms derived from the water solvent are highlighted in blue.

Following formation of T-22, Tsn15 then catalyses its conversion into tetronasin. Tsn15 could catalyse tetrahydropyran formation and oxadecalin fragmentation either via nucleophilic attack of C17 onto C13 (Mechanism 5, 6, 7 and 8) or a pericyclic rearrangement (Mechanism 9 and 10).

Depending on the mechanisms used by Tsn11 and Tsn15 to convert T-16 into tetronasin via T-22, the oxygen of the C3 keto group in tetronasin could be derived from either water (mechanisms 7 and 10), or (*2S*)-methylmalonyl-CoA (mechanisms 5, 6, 8, and 9). Feeding experiments using a precursor of ^{18}O-labelled (*2S*)-methylmalonyl-CoA previously demonstrated that the oxygen of C3 keto is derived from (*2S*)-methylmalonyl-CoA[175]. To confirm this here, the Tsn11 and Tsn15 assays were conducted in excess $H_2{}^{18}O$ and assessed for the incorporation of ^{18}O. These assays revealed no incorporation of solvent ^{18}O in either T-22 or tetronasin (**Figure 5.32**). The lack of incorporation of oxygen from the water solvent there makes mechanisms 7 and 10 unlikely, favouring Tsn11 catalysing an apparent $_{INV}$HDA reaction (analogous to the [4+2] cycloaddition catalysed by its homologue PyrE3[161]). While mechanism 8 (the other mechanism where Tsn11 catalyses a stepwise reaction), is still theoretically possible, it seems unlikely as it would require the minor form(s) of T-22 (with an open hemiacetal ring) (**Figure 5.27**) to be the true substrate(s) of Tsn15. As the oxadecalin moiety of T-22 is not present in tetronasin, the reaction catalysed by Tsn11 is latent. Discriminating between the possible mechanisms of Tsn15 (nucleophilic attack vs. pericyclic) was more difficult and required the enzyme structural studies discussed in Chapter 6. Regardless of the mechanism, it is clear that both Tsn11 and Tsn15 catalyse different reactions from their homologues in spirotetronate/spirotetramate biosynthetic pathways.

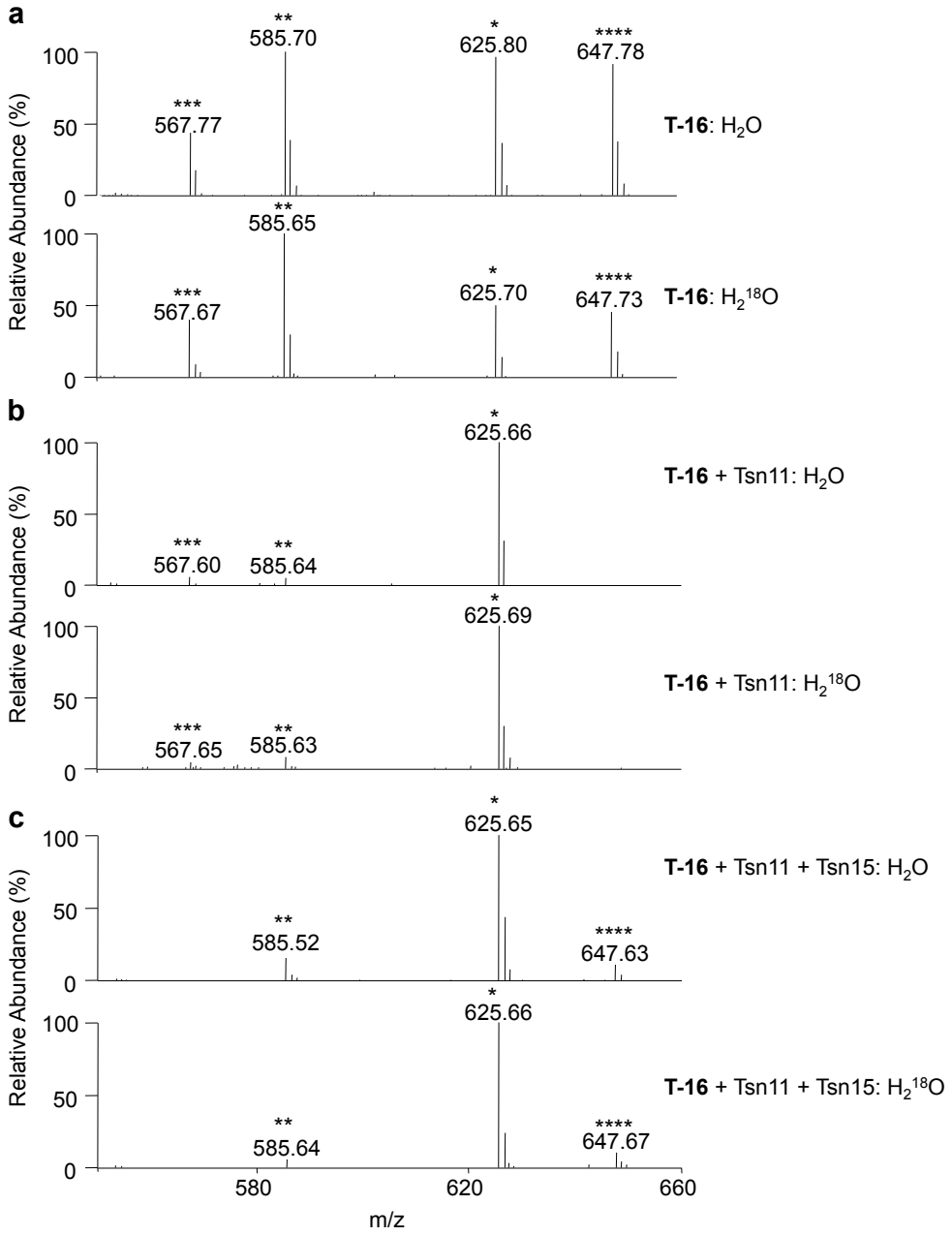


Figure 5.32 Mass spectra of the Tsn11 and Tsn15 assay performed in $H_2{}^{18}O$. a, Mass spectra of T-16 incubated without enzyme in H_2O or $H_2{}^{18}O$. **b,** Mass spectra of T-22 generated by incubating T-16 with Tsn11 in H_2O or $H_2{}^{18}O$. **c,** Mass spectra of tetronasin generated by incubating T-16 with Tsn11 and Tsn15 in H_2O or $H_2{}^{18}O$. *$[M+Na]^+$, **$[M+H-H_2O]^+$, ***$[M+H-2.H_2O]^+$, ****$[M+H+2.Na]^+$. For **b,** *$[M+Na-H_2O]^+$, **$[M+H-2.H_2O]^+$, ***$[M+H-3.H_2O]^+$.

5.3 The role of the FAD prosthetic group of Tsn11

None of the mechanisms presented in Figure 5.31 requires the bound FAD prosthetic group of Tsn11 to perform redox chemistry. As in the case of PyrE3[161,163], FAD has no

obvious role in the apparent $_{INV}$HDA catalysed by Tsn11. As discovered in Chapter 1, Tsn11 also lacks several of the residues predicted to be involved in oxidising NADPH. Nonetheless, it was important to determine experimentally whether or not the FAD prosthetic group participates in Tsn11 catalysis. A first approach to achieve this was to disrupt the interaction between Tsn11 and FAD by mutating a key FAD binding residue. The ability of Tsn11 lacking FAD to convert T-16 to T-22 could then be assayed *in vitro*. From the homology between Tsn11 and PyrE3/OxyS Asp282 was identified as likely to be an essential FAD binding residue[161,259] (**Figure 5.33**).

```
Tsn11       PPADQPVSVEEVIASTERITGRTPVVGEVQWLHRYTNTTREAENYRQGRVFVAGDAAHLH 287
PyrE3       PPADQEVGFDELRAAVARIAGVEL-DGVPGWLSRWTATSRQAERYREGRILLAGDAAHTL 265
OxyS        ADRRAPVTEEDVRAALTEVAGSDFGMHDVRWLSRLTDTSRQAERYRDGRVLLAGDACHIH 293
                  *  ::: *:  .::*         ** * * *:*:**.**:**::::****.*
```

Figure 5.33 FAD binding residue of Tsn11. Alignment of OxyS and PyrE3 with Tsn11 enabled Asp282 to be identified as a likely residue involved in binding FAD.

tsn11 D282A was successfully created using site-directed mutagenesis and cloned into a pET28a(+) plasmid for expression. Unfortunately expression of *tsn11* D282A *E. coli* only yielded insoluble (albeit no longer yellow) protein. The lack of solubility and colour of Tsn11 D282A suggested that disrupting FAD binding altered protein folding, decreasing solubility. As Tsn11 D282A could not be purified, our next approach was to prevent any redox cycling by keeping the FAD prosthetic group reduced for the duration of the assay. The *E. coli* flavin reductase protein Fre had previously been shown to selectively reduce the FAD prosthetic group of PyrE3[163], so the *fre* gene was amplified from the genome of *E. coli* and overexpressed from pET-28a(+) for Ni-affinity purification (**Appendix Figure 6**). Although the purified Fre did possess flavin reductase activity (**Appendix Figure 6**), it was unable to reduce the FAD prosthetic group of Tsn11 (**Figure 5.34a**). Notably, no reduction of FAD was detected when Tsn11 was incubated with either NADH or NADPH, indicating that its FAD prosthetic group is unable to oxidise these common redox cofactors. The chemical reducing agent sodium dithionite ($Na_2S_2O_4$) was tried next, also previously shown to be effective at reducing the FAD prosthetic group of PyrE3[163]. Incubating Tsn11 with sodium dithionite successfully reduced its FAD group, detectable both visually (yellow to colourless) and by its UV/Vis absorbance spectrum (**Figure 5.34a, Figure 5.34b**).

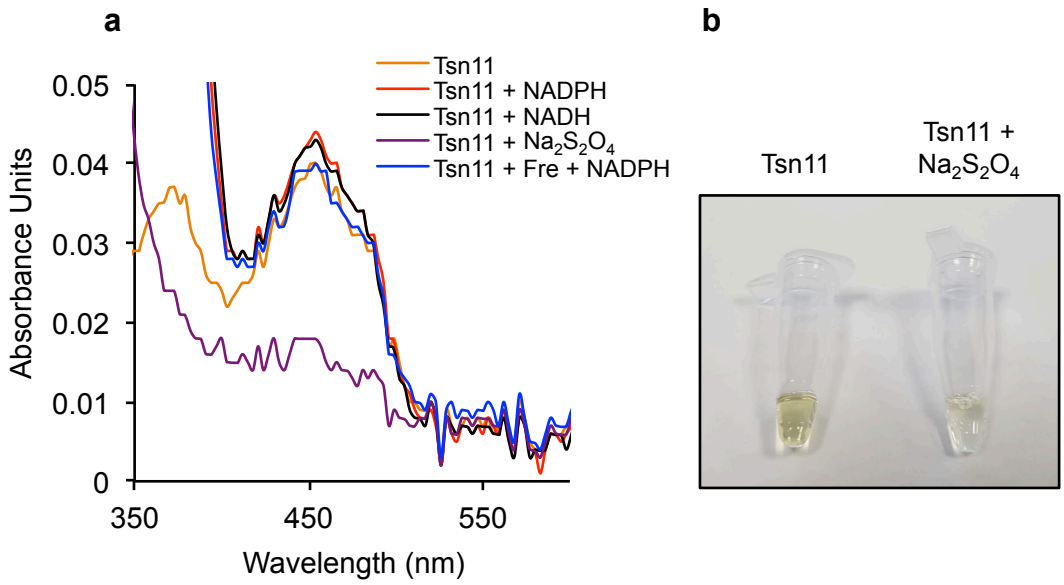


Figure 5.34 Reduction of the Tsn11 FAD prosthetic group in the presence of sodium dithionite. a, UV/Vis spectra of 100 µM Tsn11 in the presence of NADH (1 mM), NADPH (1 mM), sodium dithionite (3 mM), or Fre (50 µM) + NADPH (1 mM). Flavin reduction was only observed in the presence of sodium dithionite. **b,** bleaching of the oxidised FAD prosthetic group of Tsn11 due to reduction by sodium dithionite.

However, no inhibition of the ability of Tsn11 to convert T-16 into T-22 was detected in the presence of excess sodium dithionite (**Figure 5.35**). These experiments demonstrated that, as proposed for PyrE3[161,163], the FAD prosthetic group of Tsn11 does not participate in any redox chemistry during catalysis.

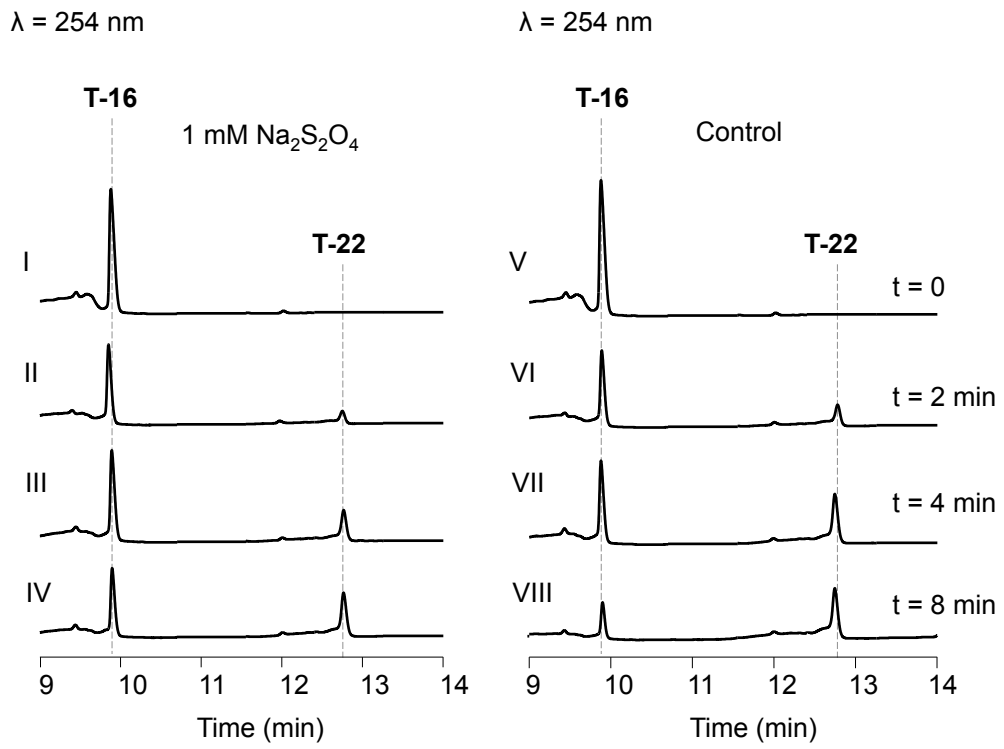


Figure 5.35 Tsn11 assay in the presence of sodium dithionite. Conversion of T-16 into T-22 by Tsn11 was monitored over time in the presence of excess of $Na_2S_2O_4$ (sodium dithionite). 100 µM of T-16 was incubated with 1 µM Tsn11 in the presence of 1 mM sodium dithionite at 30 °C. (I) t = 0 min control. (II) reaction terminated at 2 min (III) reaction terminated at 4 min (IV) reaction terminated at 8 min. Control reactions were performed in the absence of sodium dithionite (V) t = 0 min control. (VI) reaction terminated at 2 min (VII) reaction terminated at 4 min (VIII) reaction terminated at 8 min. Data are representative of three independent experiments.

5.4 Exchanging Tsn11 and Tsn15 with homologues from the *tmn* and *mad* pathways

Given the structural similarities between tetronomycin and tetromadurin with tetronasin, their cyclohexane and tetrahydropyran rings are likely to be formed by the same mechanism. The four stereocenters established in the reactions catalysed by Tsn11 and Tsn15, C4, C5, C10, and C13, all have the opposite configuration in tetronomycin. It was therefore wondered if these stereocenters could be inverted by substituting Tsn11 and/or Tsn15 for their respective *tmn* homologues, Tmn9 and Tmn8. Furthermore, as tetromadurin has the same stereochemistry as tetromadurin at these centres, attempting to substitute Tsn11 and Tsn15 for the Mad10 and Mad31, respectively, would provide insight into the substrate tolerance of the *mad* enzymes. To achieve this, the T-16 *in vitro* assays

were performed as before but Tsn11 and Tsn15 were replaced by their *mad* and *tmn* homologues. Tmn9, Tmn8, and Mad31 were readily purified as recombinant proteins from *E. coli*, as was done for Tsn11 and Tsn15 (**Figure 5.36**). Unfortunately Mad10 could not be purified as soluble protein.

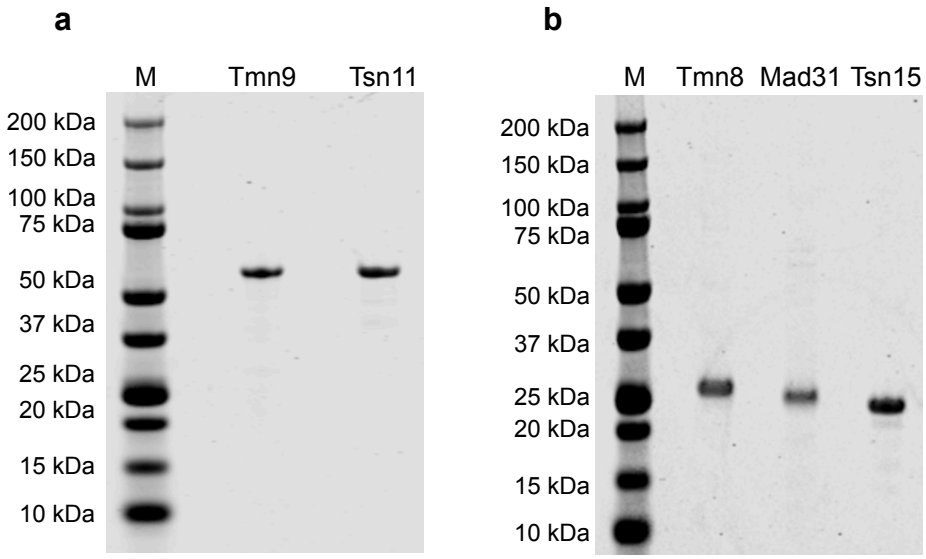


Figure 5.36 Purification of Tmn9, Tmn8, and Mad31. a, Purification of Tmn9 expressed from *E. coli* BL21 (λDE3). Marker (M); Tmn9, 56.3 kDa; Tsn11, 55.5 kDa. **b** Purification of Tmn8 and Mad31 expressed from *E. coli* BL21 (λDE3). Marker (M); Tmn8, 21.9 kDa; Mad31, 22.0 kDa; Tsn15, 24.6 kDa. Protein bands were visualised using Coomassie Blue.

In vitro assays were then performed containing T-16 and all possible combinations of the purified Tsn11 and Tsn15 homologues. The assay results demonstrated that Tmn8, Tmn9, and Mad31 have no activity with T-16 or T-22 *in vitro* (**Figure 5.37**).

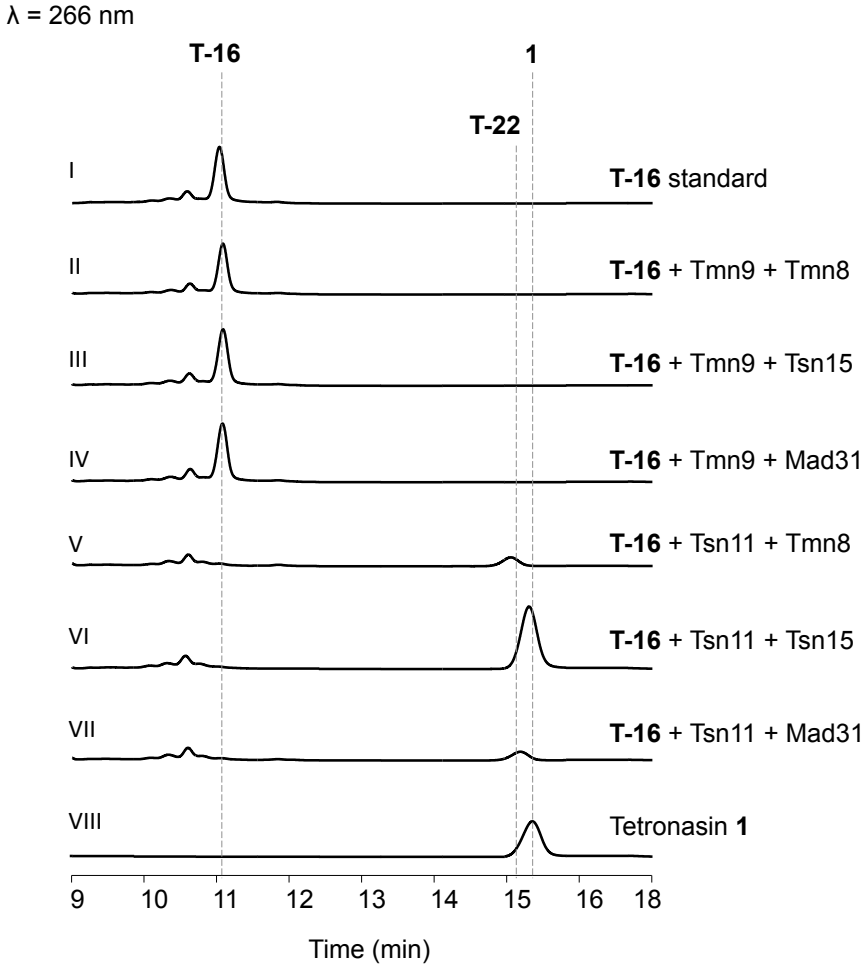


Figure 5.37 *in vitro* enzyme swapping experiments using Tsn11 and Tsn15 homologues. T-16 was incubated with Tsn11 and Tsn15 homologues before the reaction was analysed by HPLC. I) T-16 standard**.** II) T-16 incubated with Tmn9 and Tmn8. III) T-16 incubated with Tmn9 and Tsn15. IV) T-16 incubated with Tmn9 and Mad31. V) T-16 incubated with Tsn11 and Tmn8, producing T-22. VI) T-16 incubated with Tsn11 and Tsn15, producing tetronasin **1**. VII) T-16 incubated with Tsn11 and Mad31, producing T-22. VIII) Tetronasin **1** standard. Data representative of three independent experiments.

To test the substrate tolerance of *tmn* and *mad* enzymes *in vivo*, genetic cross complementation experiments in *S. longisporoflavus* ∆tsn11 and *S. longisporoflavus* ∆tsn15 were performed. Genetic complementation of the *S. longisporoflavus* ∆tsn11 strain using pIB139-*mad10* did not rescue tetronasin production (**Figure 5.38**). Likewise, genetic complementation of *S. longisporoflavus* ∆tsn11 using pIB139-tmn9 and of the *S. longisporoflavus* ∆tsn15 strain using pIB139-*tmn8* or pIB139-*mad31* failed to restore tetronasin production or produce any novel metabolites. Despite the structural similarities between tetronasin, tetronomycin, and tetromadurin, the *tmn* and *mad* Tsn11 and Tsn15 homologues have no detectable activity with the tetronasin intermediates T-16 and T-22.

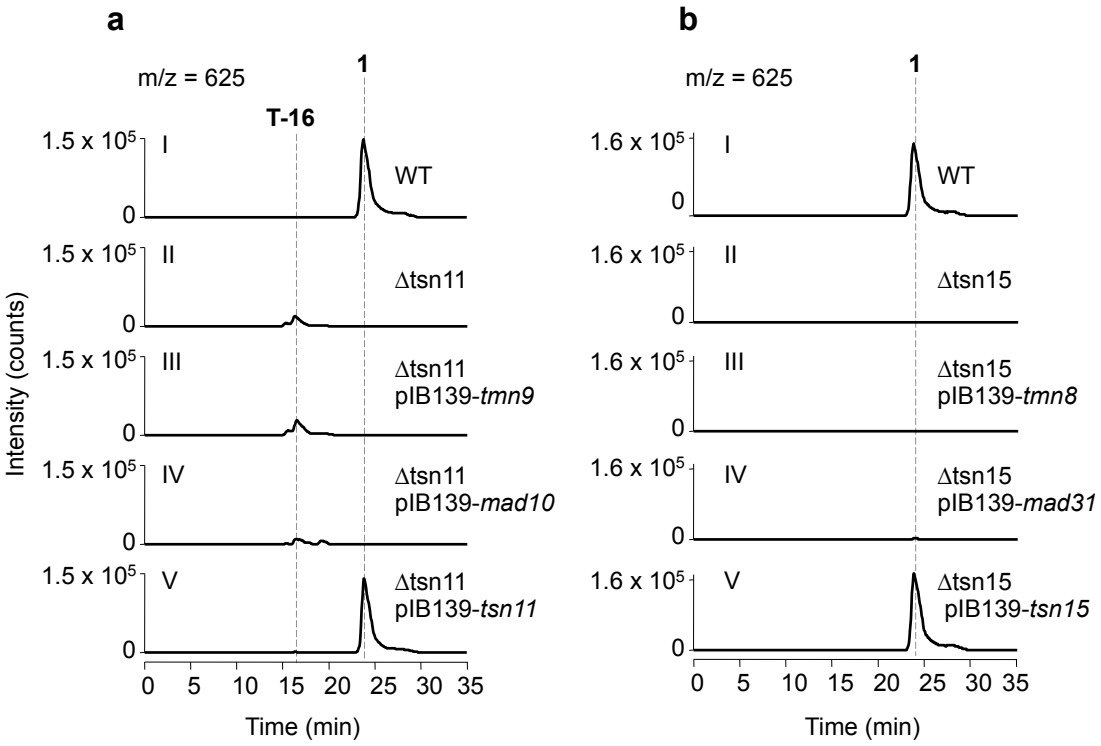


Figure 5.38 ***in vivo*** **cross-complementation using Tsn11 and Tsn15 homologues. a,** Genetic complementation of the *S. longisporoflavus* Δtsn11 mutant. I) Production of tetronasin **1** by *S. longisporoflavus* WT. II) Production of T-16 by *S. longisporoflavus* Δtsn11. III) Production of T-16 by *S. longisporoflavus* Δtsn11 pIB139-*tmn9*. IV) Production of T-16 by *S. longisporoflavus* Δtsn11 pIB139-*mad10*. V) Production of tetronasin **1** by *S. longisporoflavus* Δtsn11 pIB139: *tsn11*. **b,** Genetic complementation of the *S. longisporoflavus* Δtsn15 mutant. I) Production of tetronasin **1** by *S. longisporoflavus* WT. II) Lack of tetronasin production by the *S. longisporoflavus* Δtsn15 mutant. III) Lack of tetronasin production by the *S. longisporoflavus* Δtsn15 pIB139: *tmn8* mutant. IV) Lack of tetronasin production by the *S. longisporoflavus* Δtsn15 pIB139: *mad31* mutant. V) Production of tetronasin **1** by *S. longisporoflavus* Δtsn15 pIB139: *tsn15.* Data representative of three independent experiments.

5.5 Order and timing of ring formation

The final experiment performed in this chapter concerned the timing of cyclohexane and tetrahydropyran formation. Namely, it was wondered if Tsn11 and Tsn15 act while the tetronasin intermediate was still PKS bound, or after chain release has occurred. A chemical technique recently developed by collaborator Dr Manuela Tosin (University of Warwick) has enabled PKS-bound intermediates to be prematurely offloaded *in vivo*[281–283], enabling analysis like never before. The technique uses compounds such as methyl 6-decanamido-2-fluoro-3-oxohexanoate (M623), which compete with (2*S*)-methylmalonyl-ACP to accept the polyketide acyl chain during a PKS extension reaction[281,282]. M623 is initially an ester to increase its lipophilicity and facilitate uptake across the cell

membrane[283]. Once within the cell, ester hydrolysis occurs (catalysed by intrinsic esterases) and the activated M623 can begin offloading polyketide intermediates[283] (**Figure 5.39**). The identity of these intermediates can often be inferred using high-resolution mass spectrometry.

All M623 experiments were performed by Robert Jenkins and Dr. Manuela Tosin (University of Warwick (UK)) using cultures of *S. longisporoflavus*, *S. longisporoflavus* Δtsn11, and *S. longisporoflavus* Δtsn15. The growth medium of all three strains was supplemented with M623, followed by extraction using organic solvents and analysis by HRMS. It was hypothesised that if cyclohexane and/or tetrahydropyran rings form during chain extension on the PKS, then disrupting formation of these rings would affect the profile of intermediates offloaded by M623[196] compared to the wild type.

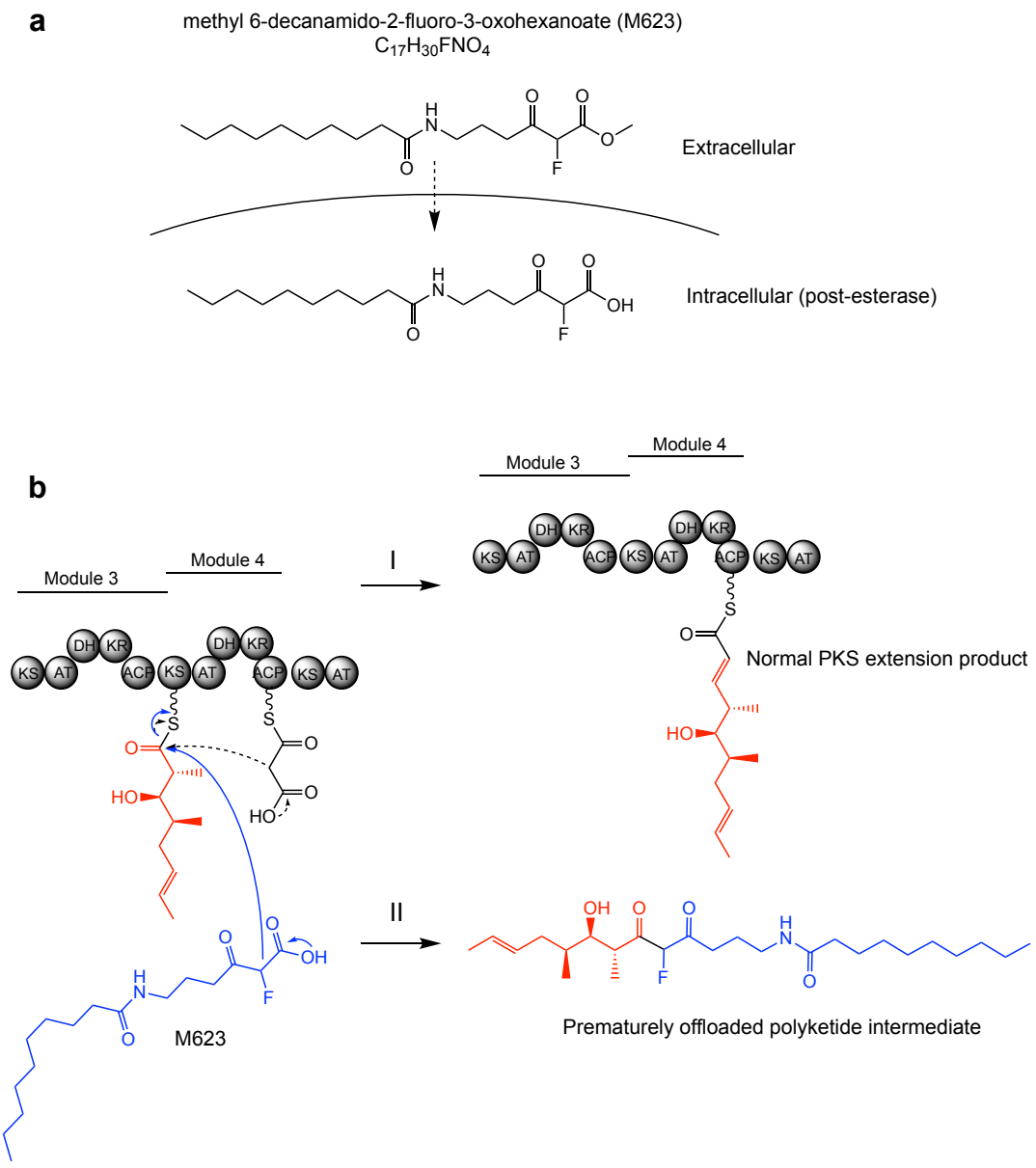


Figure 5.39 Offloading of tetronasin intermediates by methyl 6-decanamido-2-fluoro-3-oxohexanoate (M623). a, M623 crosses the cell membrane in the ester form. Upon entering the cell, intracellular esterase enzymes catalyse ester hydrolysis to produce the acid form. **b,** Mechanism of polyketide intermediate offloading by M623. I) Standard extension cycle. II) Alternative M623 extension product released from the PKS.

However, the results from these experiments revealed no major differences in the profiles of offloaded intermediates between *S. longisporoflavus*, *S. longisporoflavus* Δtsn11, and *S. longisporoflavus* Δtsn15, indicating cyclohexane and tetrahydropyran formation likely occurs post-PKS. When grown with M623 all three strains produced an undecaketide tetronasin intermediate containing the tetrahydrofuran ring (**Figure 5.40**), indicating that this ring likely forms while the tetronasin intermediate is still bound to the PKS. The

undecaketide tetronasin intermediate also contains the C30 hydroxylation and C25 methoxy group, suggesting these modifications also occur prior to chain release.

Taken together, the results from this chapter suggest that the first ring to form is the tetrahydrofuran ring, which occurs during chain extension. The next ring formed is the tetronate, which also results in release of T-16 from the PKS. Tsn11 then converts T-16 to T-22, in which the cyclohexane ring has formed (albeit as part of a larger oxadecalin). Finally, Tsn15 catalyses formation of the tetrahydropyran and fragments the oxadecalin, producing tetronasin (**Figure 5.41**). These results demonstrate that Tsn11 and Tsn15 have evolved to catalyse novel reactions from their homologues in spirotetronate/spirotetramate biosynthesis, setting the stage for the detailed structural analysis of these proteins, and closer examination of their mechanism, as described in the next chapter.

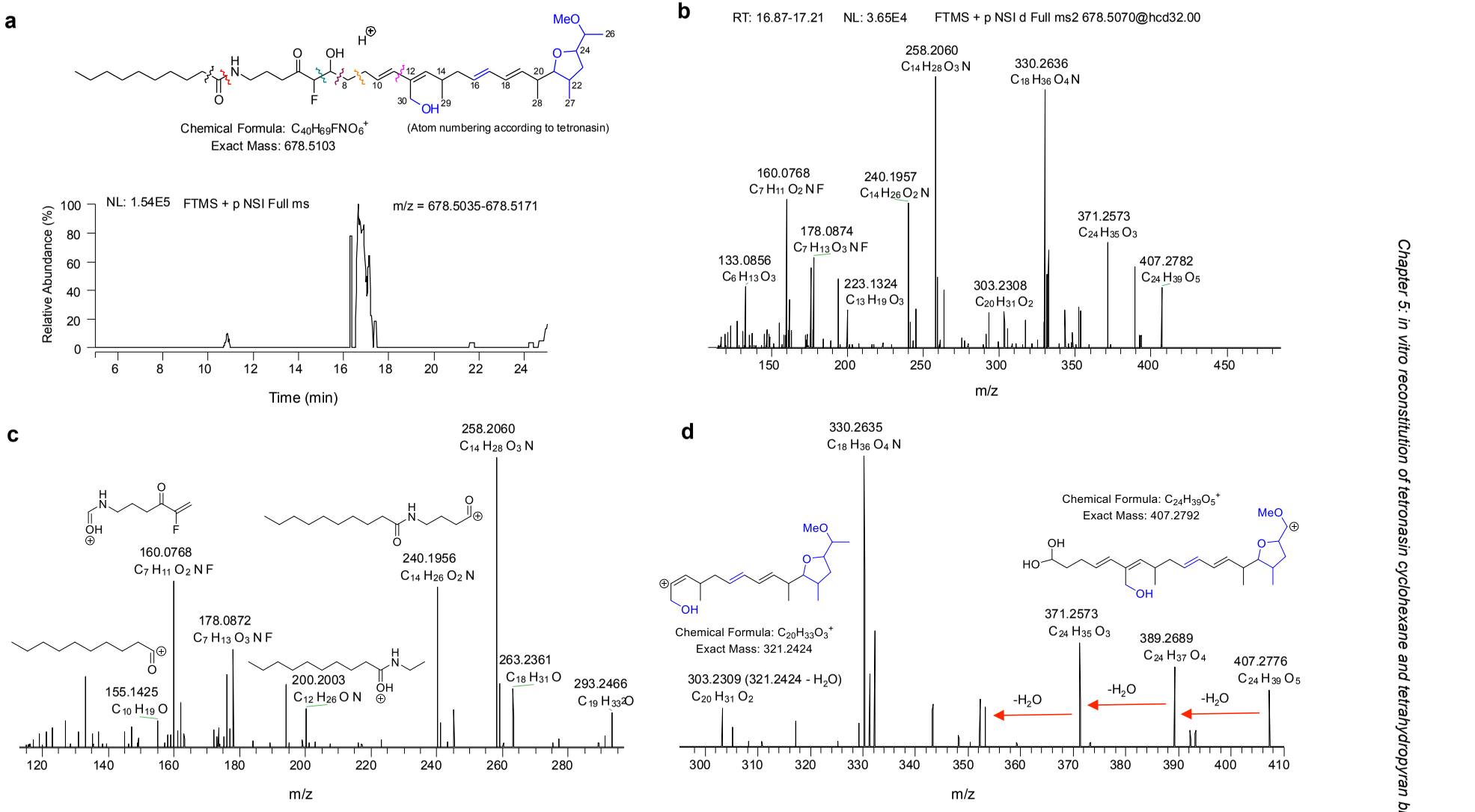

a
Chemical Formula: C40H69FNO6+
Exact Mass: 678.5103
(Atom numbering according to tetronasin)
NL: 1.54E5 FTMS + p NSI Full ms
m/z = 678.5035-678.5171
Relative Abundance (%)
Time (min)
b
RT: 16.87-17.21 NL: 3.65E4 FTMS + p NSI d Full ms2 678.5070@hcd32.00
258.2060 C14 H28 O3 N
330.2636 C18 H36 O4 N
160.0768 C7 H11 O2 N F
240.1957 C14 H26 O2 N
178.0874 C7 H13 O3 N F
371.2573 C24 H35 O3
133.0856 C6 H13 O3
223.1324 C13 H19 O3
303.2308 C20 H31 O2
407.2782 C24 H39 O5
m/z
c
258.2060 C14 H28 O3 N
160.0768 C7 H11 O2 N F
240.1956 C14 H26 O2 N
178.0872 C7 H13 O3 N F
263.2361 C18 H31 O
293.2466 C19 H33 O
155.1425 C10 H19 O
200.2003 C12 H26 O N
m/z
d
330.2635 C18 H36 O4 N
Chemical Formula: C24H39O5+
Exact Mass: 407.2792
Chemical Formula: C20H33O3+
Exact Mass: 321.2424
303.2309 (321.2424 - H2O)
C20 H31 O2
371.2573 C24 H35 O3
389.2689 C24 H37 O4
407.2776 C24 H39 O5
-H2O
-H2O
-H2O
m/z

Figure 5.40 LC-HRMS analyses of *S. longisporoflavus* WT, Δtsn11, and Δtsn15 grown with M623. Figure made by Dr. Manuela Tosin. **a,** Extracted ion chromatogram (EIC) for a putative intercepted undecaketide ($[M+H]^+$ adduct) from the *S. longisporoflavus* Δtsn15 mutant that has undergone dehydration, oxidation, cyclisation and methylation (highlighted in blue). The same putative undecaketide species was found in *S. longisporoflavus* Δtsn11 organic extracts and *S. longisporoflavus* wild type. The key MS^2 fragments are indicated on the structure. Full spectrum $HRMS^2$ analyses of the putative undecaketide. **c,** $HRMS^2$ analyses expansion showing fragments belonging to M623. $HRMS^2$ analysis showing fragments belonging to the polyketide portion of the molecule. The species was absent when the three strains were grown in the absence of M623.

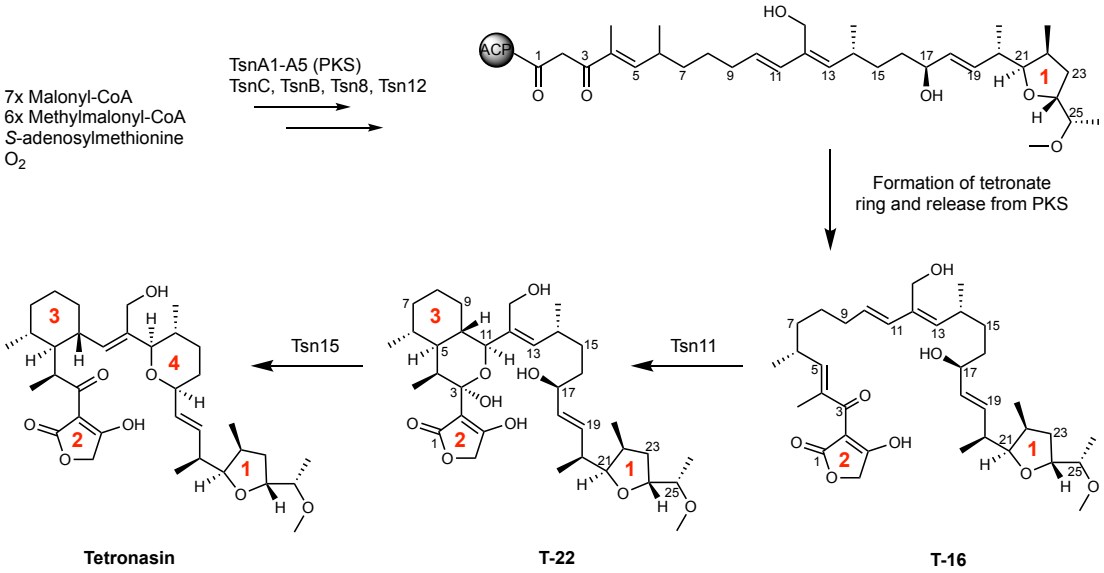


Figure 5.41 Proposed order of ring formation in tetronasin biosynthesis. The tetrahydrofuran ring forms first likely when the polyketide is still attached to the PKS. Tetronate formation is concomitant with chain release, so this is the next ring to form. Tsn11 then catalyses cyclohexane formation within a larger oxadecalin moiety. Finally, Tsn15 catalyses tetrahydropyran formation, the last ring to form, and fragments the oxadecalin to form tetronasin.

5.6 Chapter 5 Discussion

The work described in this chapter is a significant contribution to our understanding of polyether tetronate biosynthesis. Two late-stage biosynthetic intermediates of tetronasin were isolated and shown to be competent substrates of the [4+2] cyclase homologues Tsn11 and Tsn15, enabling *in vitro* reconstitution of cyclohexane and tetrahydropyran formation. From an open form of tetronasin lacking the cyclohexane and tetrahydropyran rings (T-16), Tsn11 closes the cyclohexane by producing an oxadecalin hemiacetal (T-22). Tsn15 can then catalyse tetrahydropyran formation and fragmentation of the oxadecalin, producing tetronasin.

To produce sufficient T-16 for NMR a 100 12 x 12 cm agar plates containing tsn-medium-B were grown for seven days. Though labour intensive, the upscaling of *S. longisporoflavus* ∆tsn11 required to purify sufficient T-16 for NMR was relatively minor. For a sobering comparison, to isolate the precyclised pyrroindomycin from *S. rugosporus* ∆pyrE3 (a homologue of Tsn11) the authors grew a 120 L of culture to purify only 13.4 mg of the compound[284].

Solving the structure of T-16 revealed it as an open form of tetronasin in which the cyclohexane and tetrahydropyran rings were unformed. The structure of T-16 gave several key insights into tetronasin biosynthesis. Foremost, given that T-16 accumulated as a result of deleting *tsn11*, it provided the first evidence that cyclohexane and tetrahydropyran formation is enzymatically controlled. Secondly, it demonstrated that tetronate formation, tetrahydrofuran formation, C25 O-methylation, and C30 hydroxylation are not dependent on the cyclohexane or tetrahydropyran rings. The missing ^{13}C NMR peaks for the tetronate portion of T-16 indicate that rotation around the C2-C3 double bond occurs much more readily than it does in T-22. The different NMR solvents used for these compounds could account for this. The d-chloroform used for T-16 (unlike the d-MeOH used for T-22) may have been slightly acidic, increasing rotation around the C2-C3 bond. Changing the T-16 NMR solvent to deuterated MeOH could therefore resolve this issue in the future.

The structure of T-16 also provided compelling evidence that M-17, the tetromadurin metabolite produced by *A. verrucosispora* ∆mad10 is the equivalent open form of

tetromadurin lacking cyclohexane and tetrahydropyran rings. As discussed in **Section 4.32**, M-17 is missing one of the cytochrome P450-catalysed hydroxylations. That T-16 is hydroxylated at C30 strongly suggests that M-17 is also hydroxylated at the equivalent residue (C36). The missing hydroxyl group of M-17 is therefore likely that on C38 and its insertion is dependent on cyclohexane formation (**Figure 5.42**). Cyclohexane and tetrahydropyran formation are therefore not necessarily the final two steps in polyether tetronate biosynthesis.

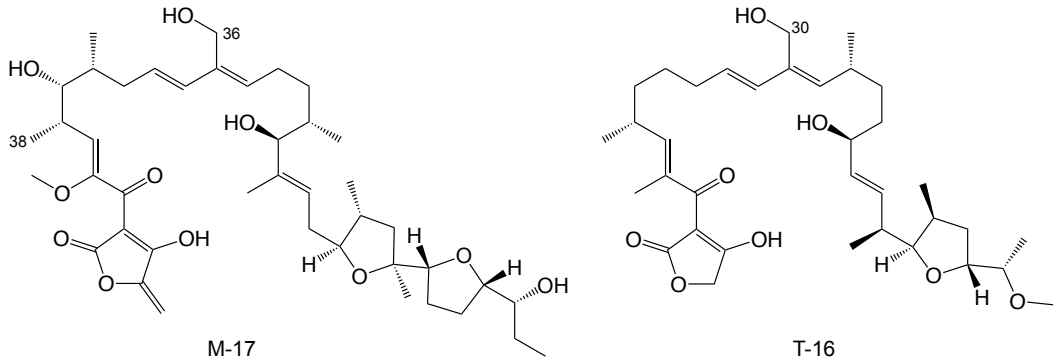


Figure 5.42 Proposed structure of M-17 based of the structure of T-16. The hydroxylation of C30 in T-16 suggests it is not dependent on the cyclohexane and tetrahydropyran rings being present. The equivalent carbon in M-17, C36, is therefore also likely to be hydroxylated. Hydroxylation of C38 is proposed to occur after cyclohexane and tetrahydropyran ring formation.

Following the purification of structural characterisation of T-16, its competency as a substrate for Tsn11 and/or Tsn15 was tested. Tsn11 and Tsn15 were readily purified using *E. coli* as an expression host. Like PyrE3, Tsn11 purified with a bound FAD prosthetic group (**Figure 5.8, Figure 5.10**)[161]. Incubating T-16 with Tsn11 resulted in conversion into a new metabolite, T-22. Tsn15 could then convert T-22 into tetronasin. No conversion occurred when T-16 was incubated with Tsn15 alone. Tsn11 and Tsn15 therefore act sequentially in an enzymatic cascade to form the cyclohexane and tetrahydropyran rings of tetronasin. Like PyrI4, boiling Tsn15 did not abolish its activity, further demonstrating the stability of this class of cyclase proteins[159]. While EDTA did not affect the activity of either Tsn15 or Tsn11, it is still unknown whether sodium ions, or another non-chelatable metal[175], have a role. That T-16 appears to be a less effective ionophore than tetronasin provides some evidence against this proposal, but it cannot be ruled out. Intriguingly, despite Tsn11 and Tsn15 catalysing different reactions from their homologues PyrE3 and PyrI4, respectively, in pyrroindomycin biosynthesis, the order in which they act is the

same[161]. Tetronasin may have therefore evolved from a biosynthetic pathway that initially produced a versipelostatin/chlorothricin-like spirotetronate, or *visa versa*. Either way, it is remarkable that, within the tetronate natural product family, PyrE3 and VstJ homologues have evolved to catalyse markedly different types of reactions in the polyether tetronates compared to the spirotetronates.

To characterise the structure of T-22, the Tsn11 reaction with T-16 was scaled up. The *in vitro* route was selected because the *S. longisporoflavus* Δtsn15 mutant did not produce T-22, but instead a likely breakdown product of it (T-21). Through scaling up the Tsn11 reaction, 3.3 mg of purified T-22 was isolated and its structure was solved. Unexpectedly, T-22 was found to contain a closed cyclohexane ring within a larger oxadecalin moiety. The tetrahydropyran ring was still unformed in T-22, indicating that it forms after the cyclohexane ring. The formation of the oxadecalin moiety by Tsn11 is strikingly reminiscent to dialkyldecalin formation catalysed by its homologue PyrE3 in pyrroindomycin biosynthesis[161]. Notable, as the oxadecalin moiety is absent form tetronasin, the reaction of Tsn11 is latent. To convert T-22 into tetronasin, T-22 must undergo tetrahydropyran formation, dehydration of the hemiacetal hydroxyl group, and oxadecalin fragmentation. Remarkably, all three of these transformations occurred when T-22 was incubated with Tsn15 (**Figure 5.11, Appendix Figure 4**). The latent ring formation catalysed by Tsn11 and the concomitant ring formation/ring opening reactions catalysed by Tsn15 is an unusual and unexpected mechanism of cyclohexane and tetrahydropyran formation. Some parallels can be drawn to the epoxidase/epoxide hydrolase enzymes involved in tetrahydrofuran ring formation (where the epoxide ring forms and then fragments)[193,194]. Otherwise, such a reaction is unprecedented.

Six mechanisms (mechanisms 5-10) were proposed for the conversion of T-16 to tetronasin via T-22, catalysed by Tsn11 and Tsn15 (**Figure 5.31**). All of these mechanisms had parallels to mechanism 3, the reaction scheme proposed prior to knowing the identity of T-22 (**Figure 5.17**). In Mechanisms 5, 6, 9 and 10 Tsn11 converts T-16 into an oxadecalin via an apparent $_{INV}$HDA, whereas in mechanisms 7 and 8 oxadecalin formation occurs via a stepwise mechanism. Following oxadecalin formation, Tsn15-catalysed tetrahydropyran formation could either occur nucleophilic attack of the C17 hydroxyl onto C13 (as included in mechanisms 5, 6,7 and 8) or a pericyclic rearrangement (mechanisms 9 and 10). Converting T-16 into tetronasin using Tsn11 and Tsn15 in the presence of

excess $H_2{}^{18}O$ showed no incorporation of the oxygen isotope (**Figure 5.32**), consistent with previous isotopic feeding studies[175]. Mechanisms where oxygen derived from water is incorporated into tetronasin could therefore be ruled out (Mechanisms 7 and 10), favouring Tsn11 catalysing an $_{INV}$HDA. Enzymes catalysing apparent $_{INV}$HDA reactions have been described in other natural product pathways, such as LepI in leporin A biosynthesis[280] and IccD in ilicicolin biosynthesis[285]. The reaction catalysed by LepI is the most similar to Tsn11, also catalysing an inverse-electron-demand hetero (oxygen) Diels-Alder reaction to form an oxadecalin (**Figure 5.43**)[280]. Despite this, Tsn11 and LepI are unrelated, with LepI resembling a SAM-dependent methyltransferase. The mechanism of Tsn15 is explored more in the next chapter.

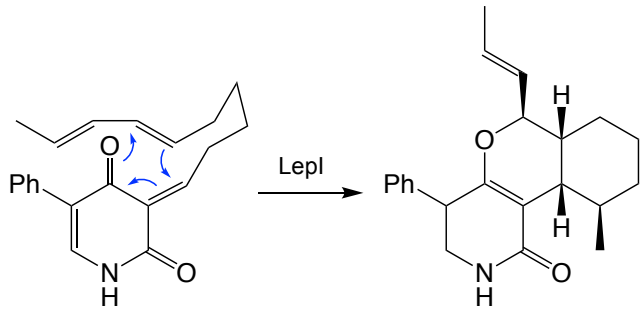


Figure 5.43 $_{INV}$HDA reaction catalysed by LepI. LepI catalyses an apparent $_{INV}$HDA reaction to form the dihydropyran core of leporin A[280]. Ph = phenol group.

Like PyrE3, the FAD prosthetic group of Tsn11 was found to not be essential for its activity[161]. Reducing the FAD group of Tsn11 using sodium dithionite did not inhibit its activity, indicating that FAD is not used for any redox chemistry. Aligning Tsn11 with PyrE3 and OxyS enabled a putative FAD binding residue (D282) to be identified. Mutating D282 to alanine prevented its soluble expression in *E. coli*. The loss of Tsn11 D282A solubility suggests that this residue is critical for the structure of Tsn11, likely though its hydrogen bonding interaction with FAD. Loss of this interaction results in decreased solubility, likely due to protein misfolding. Changes in the PyrE3 tertiary structure were observed upon mutating key FAD binding residues, confirming its important role in proper protein folding[161]. Interestingly, Tsn11 and PyrE3-family proteins are not the only [4+2] cyclase enzymes to bind FAD. The FAD-binding solanapyrone synthase, Sol5, oxidises the primary alcohol group of prosolanapyrone II to produce prosolanapyrone III. In addition to this, Sol5 also accelerates an *exo* intramolecular [4+2] cycloaddition to convert prosolanapyrone III into the *cis* decalin Solanapyrone A[137,141,142] (**Figure 5.44**). The dual activity of Sol5 sheds

light on the possible evolutionary origin of FAD-binding [4+2] cyclases. Exclusive FAD-binding [4+2] cyclase enzymes such as PyrE3 and Tsn11 may have initially been monooxygenases that developed a promiscuous [4+2] cyclase side activity like Sol5. The evolution of such promiscuous activities in enzymes is tolerated as they often do not affect the native function of the enzyme[286]. Over evolutionary time the monooxygenase activity could have been selected against until the [4+2] cyclase activity became the dominant activity. Whether FAD-binding proteins are more predisposed than others to evolve [4+2] cyclase activity requires further investigation.

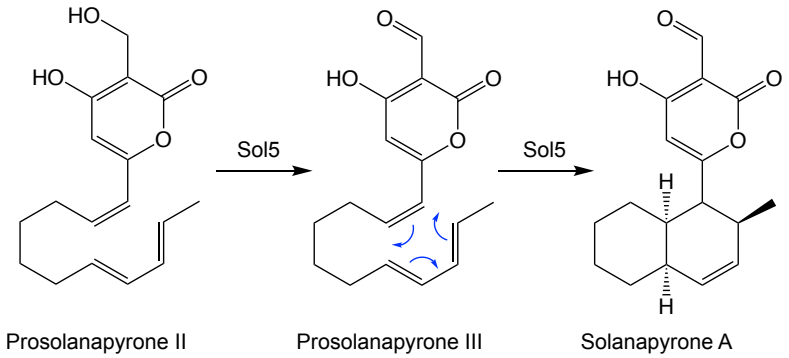


Figure 5.44 *cis*-decalin formation catalysed by the [4+2] cyclase Sol5. The FAD-binding Sol5 oxidises the primary alcohol of prosolanapyrone II, producing prosolanapyrone III. Sol5 also accelerates an *exo* Diels-Alder reaction to form a *cis* decalin. (This chapter 1 figure is repeated here for clarity).

One application of Tsn11 and Tsn15 type enzymes could be the production of novel cyclohexane and tetrahydropyran-containing polyether tetronates. Of particular interest was whether the *tmn* pathway homologues, Tmn9 and Tmn8, could produce a tetronasin analogue with opposite, tetronomycin-like stereochemistry, at C4, C5, C10, and C13. Unfortunately the Tsn11 and Tsn15 homologues from the tetronomycin (Tmn9 and Tmn8) and tetromadurin (Mad10 and Mad31) biosynthesis pathways had no activity with T-16 or T-22 (**Figure 5.37, Figure 5.38**). While these experiments were unsuccessful, other [4+2] cyclases have been shown to be tolerant of alternative substrates. For example the Tsn11 and Tsn15 homologues from the chlorothricin pathway, ChlE3 and ChlL, respectively, were able to substitute for PyrE3 and PyrI4 both *in vitro* and *in vivo* to reconstitute biosynthesis of the dialkyldecalin and spirotetramate rings of pyrroindomycin[161]. The polyketide cores of pyrroindomycin and chlorothricin are highly similar to each other in terms of stereochemistry and substituents, more so than the three polyether tetronates, likely explaining why the enzymes could be swapped. Inversion of product stereochemistry has been successfully performed by gene swapping in different systems. Fsa2 is a [4+2]

cyclase responsible for *trans* decalin formation in equisetin biosynthesis[287]. When the *fsa2* gene was inserted in place of *phm7,* a homologous gene from phomasetin biosynthesis, a phomasetin homologue was produced with equisetin-like decalin stereochemistry[288]. In the case of tetronasin and tetronomycin, it may be that the stereochemistry of C17 hydroxyl (determined by a KR domain) sterically prohibits alternative stereochemistries at C13. Other intrinsic steric constraints of T-16 may also restrict the stereochemistry of the cyclohexane ring. This idea is supported by the total synthesis of tetronasin in which both rings spontaneously formed the correct stereochemistry from an intermediate similar to T-16[180] (**Figure 1.41**). More work is clearly needed to determine why Tmn8 and Tmn9 have no activity with T-16 or T-22 and to delineate the role Tsn11 and Tsn15 have in establishing the final stereochemistry of tetronasin.

The Tsn15 [4+2] cyclase homologue AbyU from abyssomicin biosynthesis was shown to tolerance substrate analogues, enabling the creation of novel spirotetronates[158]. There is some evidence that Tsn11 and Tsn15 may be able to tolerate substrate analogues to produce novel tetronasin compounds. Feeding *S. longisporoflavus* non-natural SNAC triketide precursors resulted in the production of novel tetronasin analogues[289] (**Figure 5.45**), indicating that Tsn11 and Tsn15 can tolerate changes to the tetrahydrofuran portion of the compound. The *S. longisporoflavus* ΔtsnC mutant (**Figure 1.43**), produced a form of tetronasin completely lacking the tetrahydropyran ring, further indicates that Tsn11 and Tsn15 do not interact with this portion of the molecule.

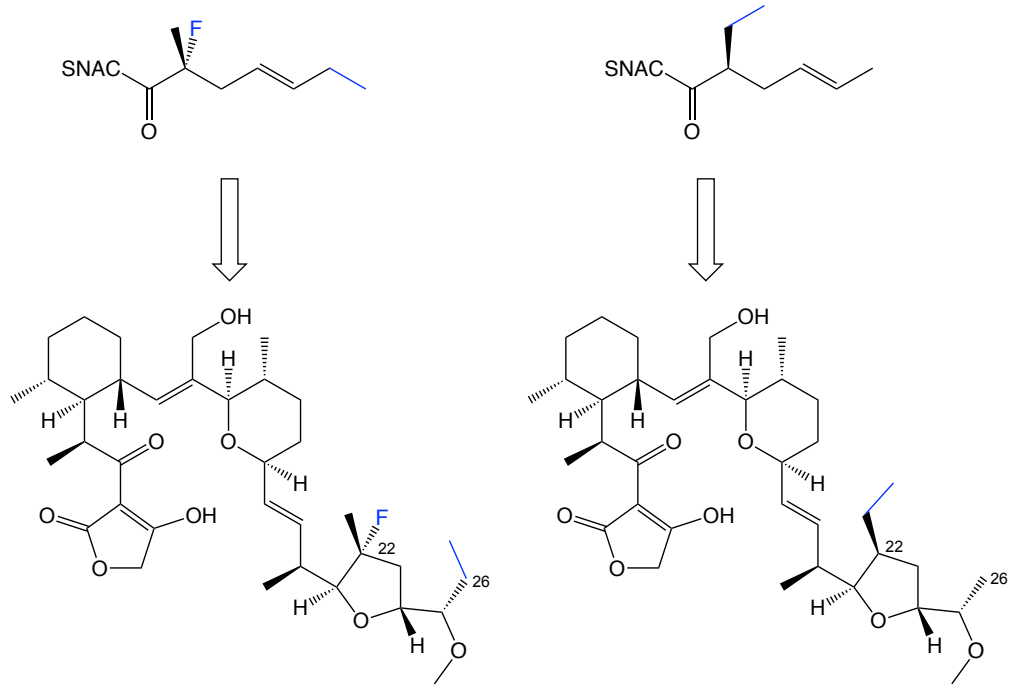


Figure 5.45 Triketide tetronasin analogue feeding experiments in *S. longisporoflavus*. Feeding tetronasin with either of the displayed SNAC-triketide analogues resulted it its incorporation into tetronasin. Figure adapted from[289]

Finally, the premature offloading experiments (**Figure 5.39**) performed by Robert Jenkins and Dr. Manuela Tosin suggested that tetrahydrofuran formation likely occurs whilst the tetronasin intermediate is PKS bound, as also predicted for the polyether lasalocid A[196]. It cannot be ruled out, however, that these enzymatic modifications actually occurred after M623 offloaded the intermediate from the PKS; *in vitro* enzymatic oxidative cyclisation assays would be required to test this. Cyclohexane and tetrahydropyran ring formation are therefore likely to be post-PKS events, making T-16 and T-22 true biosynthetic intermediates rather than shunt products. The rapid and complete turnover of T-16 and T-22 by Tsn11 and Tsn15, respectively, is further evidence they are true substrates. For comparison, deletion of the pyran synthase *salBIII* in *Streptomyces albus* resulted in a production of a precyclisation form of salinomycin lacking its tetrahydropyran ring[197]. Incubation of this precyclisation substrate with purified SalBIII for 20 h resulted in only 1% conversion to the tetrahydropyran ring, a strong hint that SalBIII acts on the ACP bound polyketide intermediate, rather than the free acid[197,290]. If the tetrahydrofuran of tetronasin does indeed from prior to chain release, then the order of ring formation would be 1) tetrahydrofuran, 2) tetronate, 3) cyclohexane, then 4) cyclohexane (**Figure 5.40**).

In summary, this chapter demonstrated that the [4+2] cyclase homologues Tsn11 and Tsn15 act in an enzymatic cascade to form the cyclohexane and tetrahydropyran rings of tetronasin. A similar cascade is likely responsible for forming the cyclohexane and tetrahydropyran rings of tetronomycin and tetromadurin (putative intermediates presented in **Figure 5.46**). Despite this, the enzymatic mechanisms underlying this transformation were unknown, leading us to perform the detailed structural and mechanistic analysis described in the following chapter.

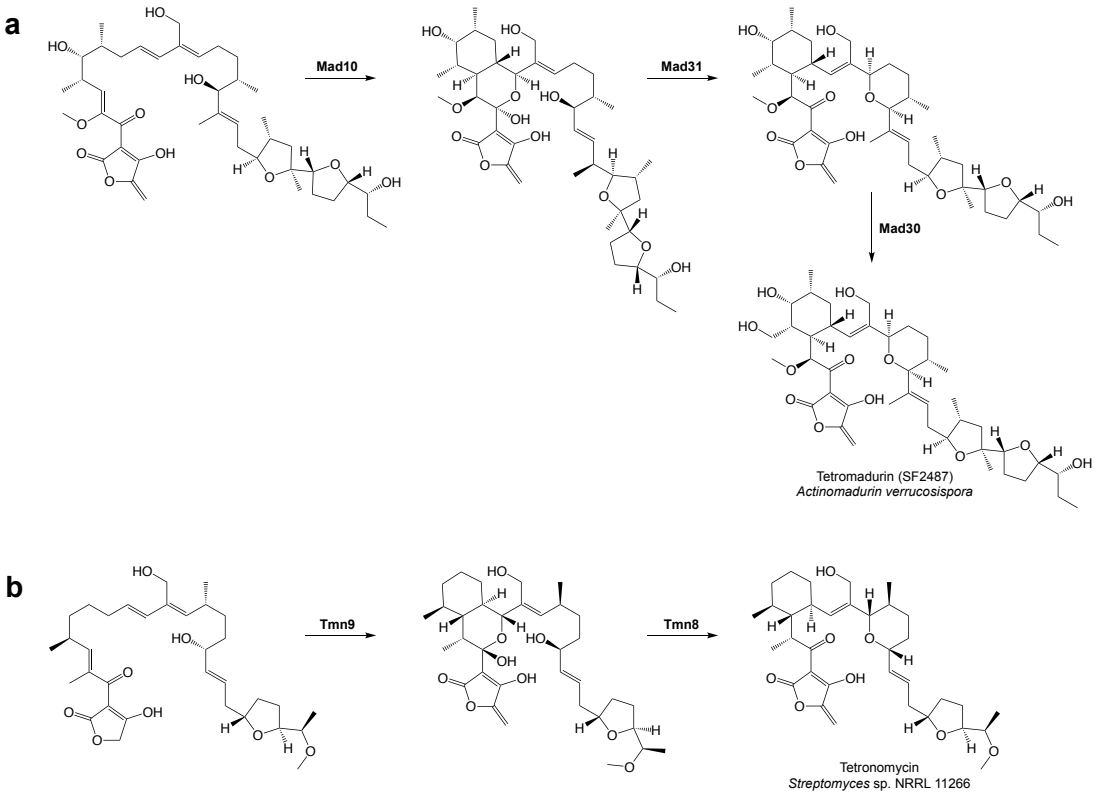


Figure 5.46 Proposed biosynthesis of cyclohexane/tetrahydropyran rings of tetronomycin and tetromadurin. Based on the structures of T-16 and T-22, the theoretical biosynthetic pathways for the biosynthesis of the cyclohexane and tetrahydropyran rings of **a,** tetromadurin and **b,** tetronomycin are presented here. The final hydroxylation of tetromadurin is predicted to occur after cyclohexane formation (see **chapter 4**).

Chapter 6: Structural studies on Tsn11 and Tsn15

A crystal structure of either Tsn11 or Tsn15 would potentially provide key insights into its active site and catalytic mechanism. However, obtaining the crystal structure of a protein is not a trivial task. Whether a protein will crystallise is highly unpredictable and numerous factors, such as pH, temperature, and salt concentration influence its success[291]. The first step towards crystallising any protein is to purify it in a single oligomeric state[292]. For crystal nucleation to occur the protein must reach a supersaturated state in an appropriate buffer[292]. Selecting the ideal protein concentration and crystallisation buffer often comes down to trial and error, much of which can now be performed by crystallisation robots. Some of the experiments described in this chapter were done by or in collaboration with Fernanda C.R Paiva and Dr. Marcio V.B Dias. In such cases the contributions are explicitly stated.

6.1 Quaternary structure of Tsn11 and Tsn15

Before attempting to crystallise Tsn11 or Tsn15, their oligomeric states in solution were determined. Knowing the oligomeric state of a protein prior to crystallisation is valuable, as the solved crystal may depict an unnatural oligomeric state. The homologues of Tsn11 and Tsn15 are typically homodimeric, though Tsn15 does have several trimeric homologues (AOC2 and PodA)[158–160,163,293]. The oligomeric state of Tsn11 and Tsn15 in solution was determined using (sedimentation velocity) analytical ultracentrifugation (AUC). AUC involves centrifuging a protein in a specialised chamber and monitoring its sedimentation rate to the bottom of the chamber[294]. The AUC data collected for Tsn11 and Tsn15 indicated that both are overwhelmingly homodimeric in solution (**Figure 6.1**), though low levels of monomer and tetramer were detected for Tsn11.

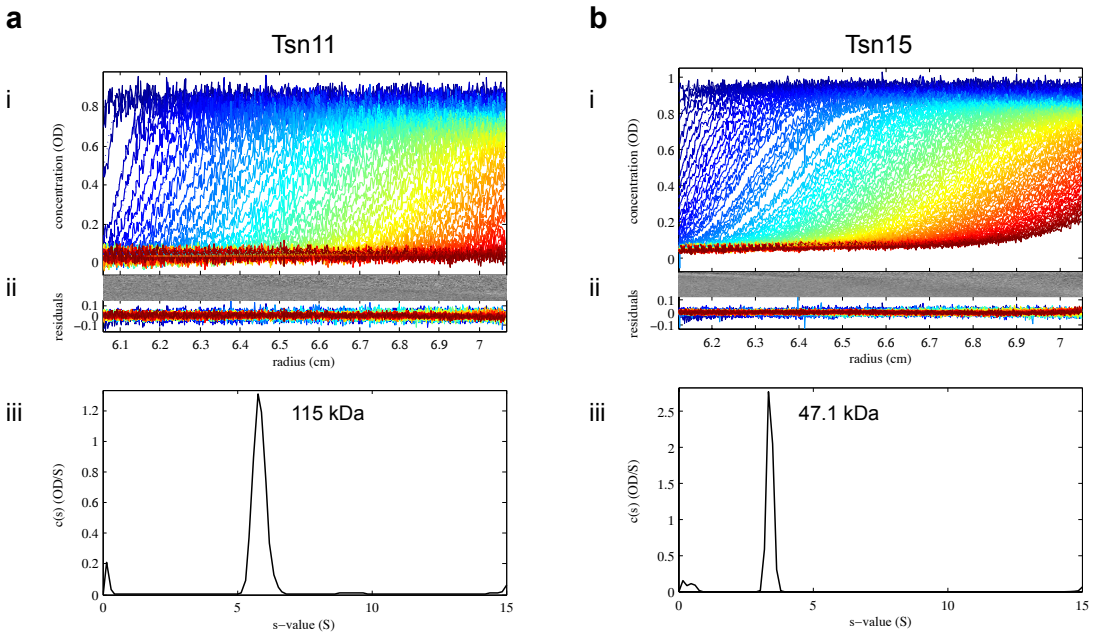


Figure 6.1 Analytical ultracentrifugation of Tsn11 and Tsn15. a, Analytical ultracentrifugation absorbance data of Tsn11 was fitted with an r.m.s deviation of 0.026, showing a mean $S_{20,w}$ (sedimentation coefficient corrected for the density and viscosity of water at 20 °C) of 5.8 and a best fit frictional ratio (f/f_0) of 1.42. The mean molecular weight calculated was 115 kDa, in good accordance with 111.0 kDa, the theoretical homodimeric mass of Tsn15 expressed with a pET28a(+) polyhistidine tag. **b,** Tsn15 analytical ultracentrifugation absorbance data was fitted with an r.m.s deviation of 0.014, showing a mean $S_{20,w}$ of 3.4 and a best fit frictional ratio (f/f_0) of 1.35. The mean molecular weight calculated was 47.1 kDa, in good accordance with 49.1 kDa, the theoretical homodimeric mass of Tsn15 expressed with a pET28a(+) polyhistidine tag. For both **a** and **b:** i) Sedimentation fringes of purified Tsn15 collected by absorbance optics. ii) Residuals from the Lamm equation solutions. iii) Continuous c(*S*) distributions of Tsn15 from the range 0-15 S. A small solvent redistribution peak is present at S = 0.

An idea from the outset of this project was that Tsn11 and Tsn15 might interact with one another, forming a ring-closing heterocomplex. While the cascade mechanism discovered in Chapter 5 made such a heterocomplex unlikely, Tsn11 and Tsn15 were nevertheless co-sedimented to confirm. As expected, the co-sedimentation data clearly demonstrated that Tsn11 and Tsn15 do not interact with each other in solution (**Figure 6.2**).

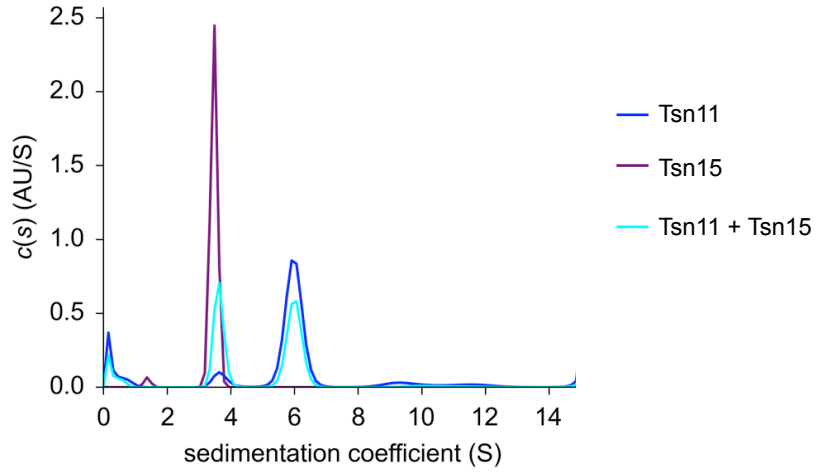


Figure 6.2 Cosedimentation of Tsn11 and Tsn15. Tsn11 and Tsn15 were mixed together and co-sedimented together in an analytical ultracentrifuge. The sedimentation of Tsn11 alone (blue), Tsn15 alone (purple), and Tsn11 and Tsn15 together (teal) were each monitored by UV absorbance compared to a buffer only control. The $S_{20,w}$ (sedimentation cooefficient corrected for the density and viscosity of water at 20 °C) of Tsn11 and Tsn15 were the same whether cosedimented alone or together, indicating that no heterocomplex formed. A small solvent redistribution peak is present at S = 0.

6.2 Purification of Tsn11 and Tsn15 for crystallisation

To produce sufficient protein for crystallisation trials, *E. coli* cultures overexpressing Tsn11 or Tsn15 were grown in 2-3L batches (typically yielding 10-15 mg/L of the overexpressed protein). Tsn11 and Tsn15 were purified using Ni-affinity chromatography as before, then further purified using preparative ÄKTA size-exclusion chromatography. As the AUC of Tsn11 and Tsn15 indicated, both proteins appeared largely monodisperse when independently passed though an ÄKTA column, though some aggregation and/or alternative states were clearly present for Tsn11 (**Figure 6.3**). The major peaks of Tsn11 and Tsn15, each corresponding to homodimer, were collected using a ÄKTA fraction collector and used for crystallisation trials.

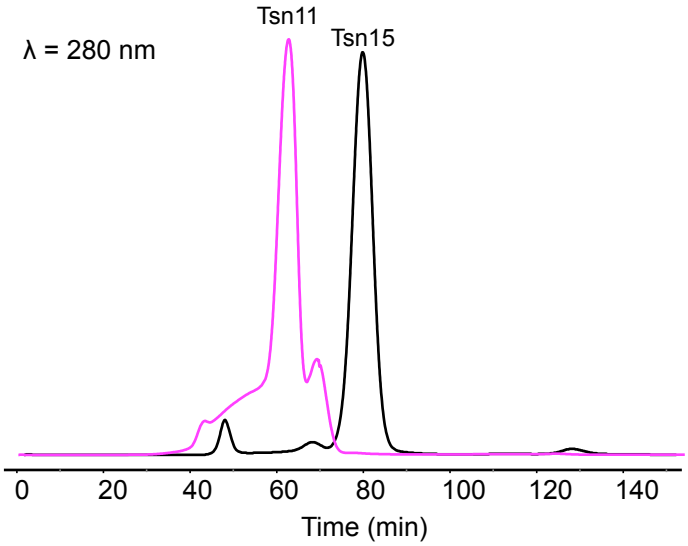
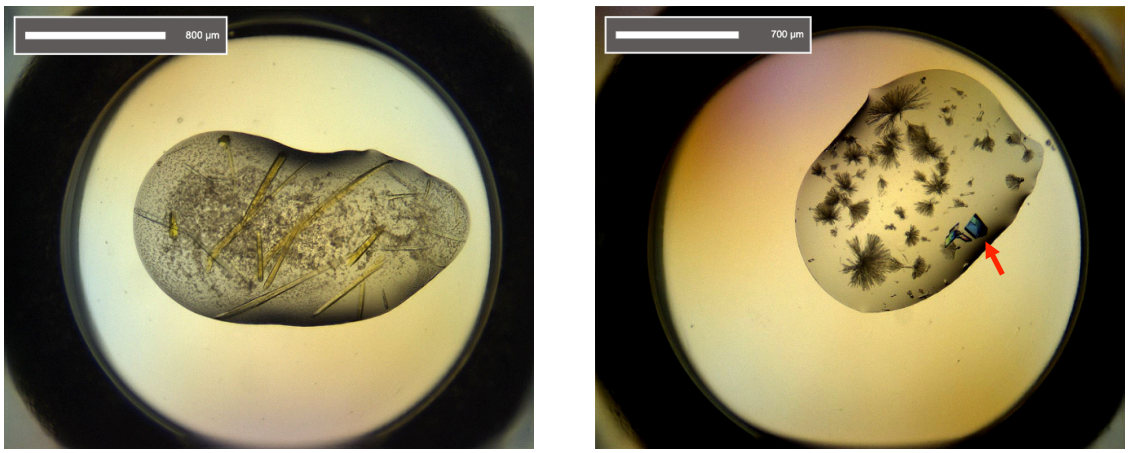


Figure 6.3 FPLC spectra of Tsn11 and Tsn15. ÄKTA gel-filtration spectra of purified Tsn11 (magenta) and Tsn15 (black).

6.3 Crystallisation of Tsn11 and Tsn15

Over 500 different crystallisation conditions (from commercial premade kits) were attempted for the crystallisation of Tsn11 and Tsn15. The Mosquito crystallisation robot (TTP Labtech) was used to mix 0.2 µL of purified protein with 0.2 µL of each different crystallisation buffer to form a sitting drop. Water slowly diffused out of these drops to concentrate the protein, eventually bringing it to a supersaturated state where crystal nucleation could potentially occur[292]. Tsn11 and Tsn15 were both successfully crystallised using this technique (**Figure 6.4**). Crystals of each protein were harvested and 'shot' at the Diamond Synchrotron (UK). Due to time constraints, only 14 crystals of Tsn11 were analysed, all of which diffracted poorly at ca. 6 Å. Optimising several conditions resulted in larger crystals growing, resulting in slightly better diffraction (ca. 4 Å). Despite this, the structure of Tsn11 could not ultimately be solved.



Tsn11 crystals Tsn15 crystals

Figure 6.4 Crystals of Tsn11 and Tsn15. Tsn11 and Tsn15 were both successfully crystallised using the sitting-drop vapour-diffusion technique. Scale bar for the Tsn11 crystal drop image = 800 µm. Scale bar for the Tsn15 crystal drop image = 700 µm. The red arrow indicates the Tsn15 crystal used to generate the x-ray diffraction pattern in **Figure 6.5**.

The Tsn15 crystals diffracted much better and from one a 2.0 Å resolution dataset was collected (**Figure 6.5**).

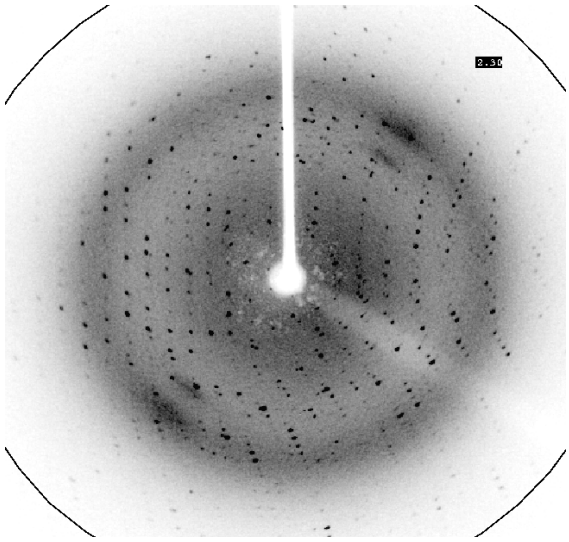


Figure 6.5 X-ray diffraction pattern of Tsn15. The X-ray diffraction pattern of Tsn15 from which a 2.0 Å dataset was collected. The Tsn15 crystal used is indicated on Figure 6.4 by an arrow.

When X-rays diffract from a crystal, only wave amplitude information is collected, while the phases of the wave remain unknown[295]. It is therefore also necessary to measure the phases (called solving the "phase problem") before the electron density of the crystal can be calculated and structural models be built. A common method of obtaining phases is called molecular replacement, where the previously solved phases of a homologous protein are used to calculate the unknown phases[295]. For molecular replacement to be successful, typically >30% identity is required between the two proteins. Unfortunately the sequence identity of the Tsn15 homologues (AbyU[158], Pyrl4[159], and AOC2[160,296]) was too low (ca. 20%) for any to be used as models for molecular replacement. An alternative method is to incorporate a heavy atom (typically selenium) into the protein structure to introduce an anomalous signal, from which the phases of the scattered waves can be calculated (called single-wavelength anomalous "SAD" phasing)[297]. Selenium can be incorporated into a protein overexpressed from *E. coli* relatively easily by first inhibiting *de novo* methionine biosynthesis and then feeding the culture D-selenomethionine (SeMet) (**Figure 6.6**).

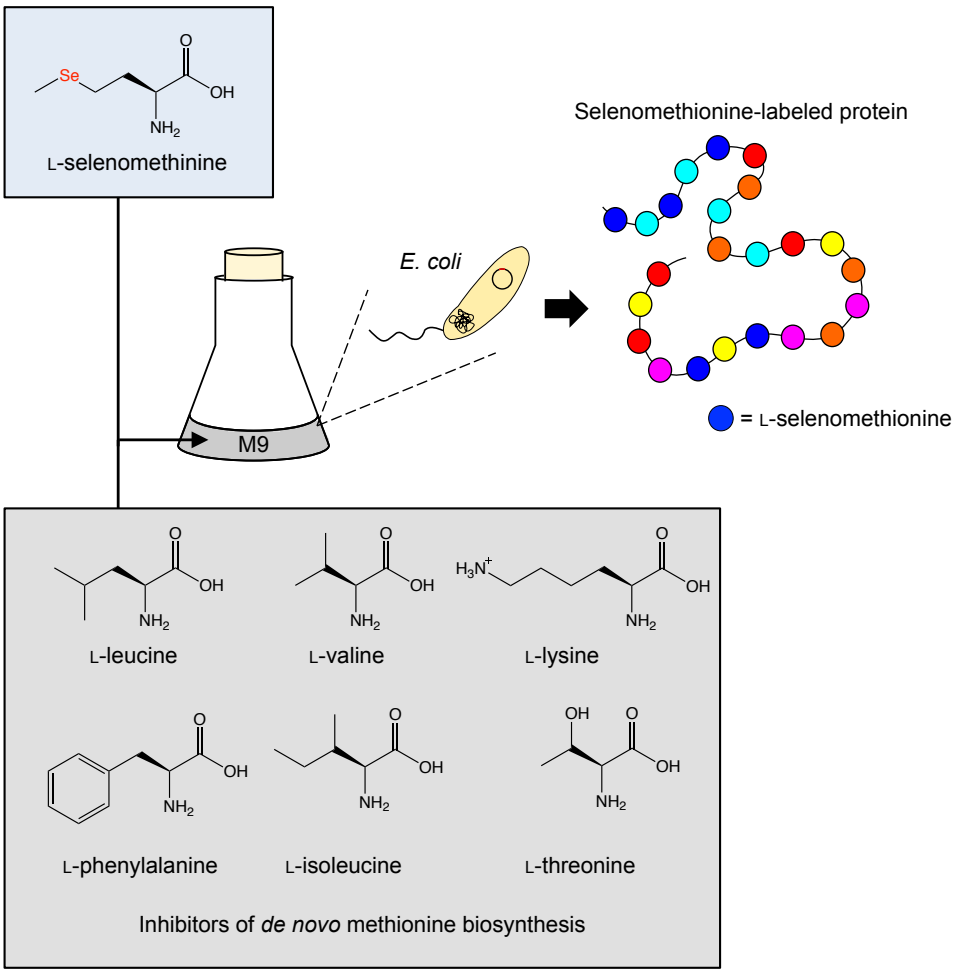


Figure 6.6 Method used to produce a selenomethionine-labelled protein. The overexpression of Tsn15 was induced from *E. coli* BL21 (λDE3) in minimal medium (M9). Prior to inducing protein expression, L-leucine, L-valine, L-lysine, L-phenylalanine, L-isoleucine, and L-threonine were added to the medium to inhibit *de novo* methionine biosynthesis. L-selenomethionine was then added and incorporated into the overexpressed protein.

When expressed from pET28a(+) with an *N*-His tag, Tsn15 has four methionine residues (excluding the *N*-terminus methionine that is typically cleaved[298]). Using the method presented in **Figure 6.6**, each of these four methionine residues was efficiently replaced by a selenomethionine, producing SeMet-Tsn15 (**Figure 6.7**). SeMet-Tsn15 was then crystallised in the same condition successfully used for Tsn15.

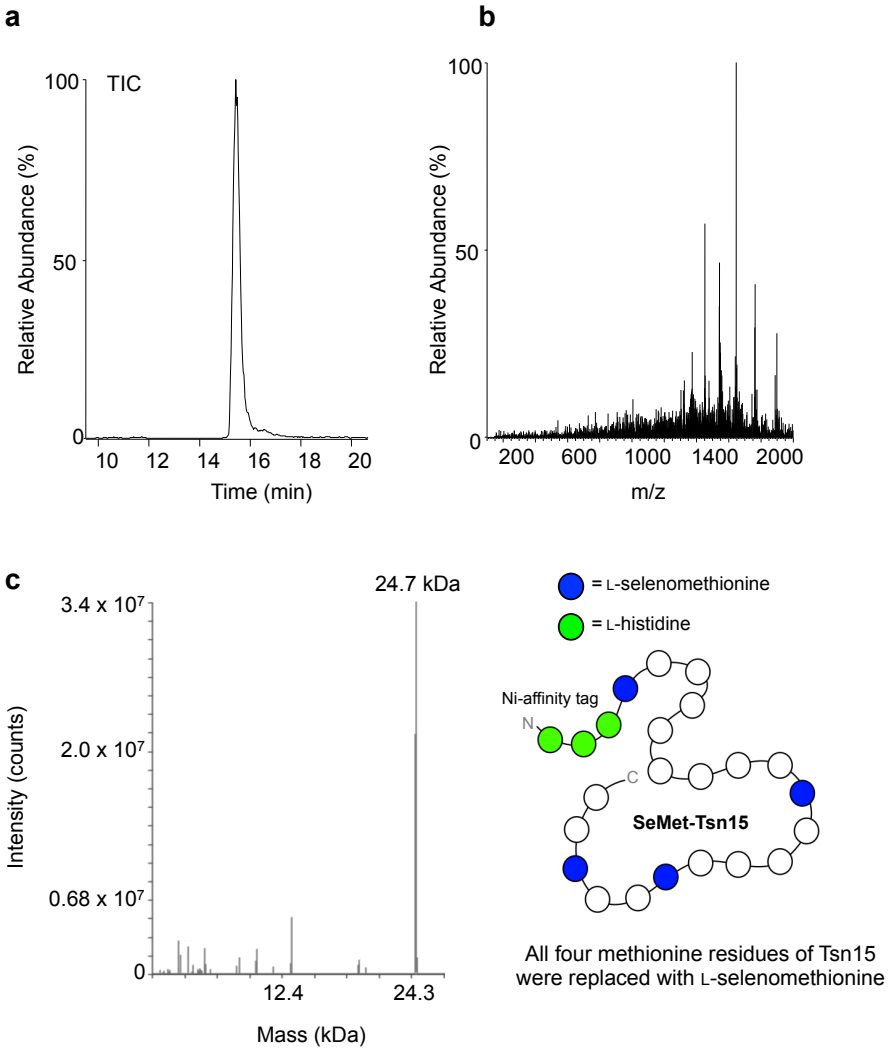


Figure 6.7 Incorporation of selenomethionine into Tsn15. a, Total ion current (TIC) spectrum of purified SeMet-Tsn15. **b,** Mass spectrum of SeMet-Tsn15. **c,** Deconvoluted mass spectrum (z=1) of SeMet-Tsn15. The unrounded mass of SeMet-Tsn15 detected was 24666.6 Da, corresponding well to 24666.9, the theoretical mass of Tsn15 where four methionines have been replaced with L-selenomethionine (and the *N*-terminal methionine has been enzymatically cleaved). The theoretical mass of unlabelled Tsn15 is 24.5 kDa.

The SeMet-Tsn15 crystal diffracted even better than the unlabelled crystals, enabling the phases to be measured and the structure of Tsn15 to be solved at 1.8 Å (**Figure 6.8**). Refinement data and statistics are presented in **Appendix Table 3**. The author did the majority of the model refinement, with some assistance from Dr. Marcio Dias.

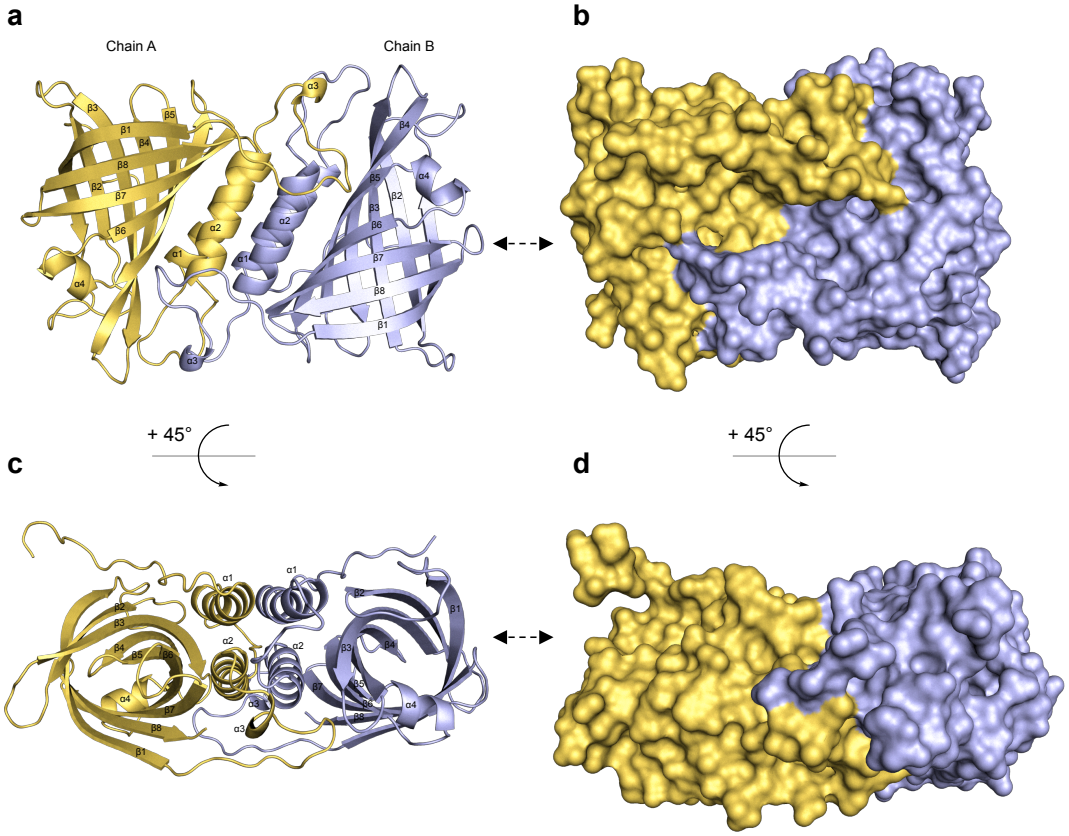


Figure 6.8 The structure of Tsn15. The structure of Tsn15 was solved to 1.8 Å using SAD phasing. **a,** Ribbon diagram of chain A (yellow) and chain B (blue) of Tsn15. **b,** Surface diagram of Tsn15. **c,** Ribbon diagram of chain A (yellow) and chain B (blue) of Tsn15 from above. **d,** Surface diagram of chain A (yellow) and chain B (blue) of Tsn15 from above.

There was a single homodimer in the asymmetric unit of Tsn15, consistent with the solution structure of Tsn15 calculated using AUC (**Figure 6.8**). The two monomers of Tsn15 are oriented in an antiparallel direction to one another. Each monomer chain is comprised of a α-helical dimerisation domain (α1-α3) and an eight-strand β-barrel (β1-β8, α4). Fernanda Paiva used the PDBsum[299] software to identify the residues, primarily through hydrophobic interactions, contributing to the dimerisation interface (primarily hydrophobic interactions) (**Appendix Figure 7**). Within each β-barrel is a largely hydrophobic pocket (**Figure 6.9**) that is likely the location of the active site. Within this cavity a polyethylene glycol (PEG) fragment (derived from the crystallisation buffer) was bound. The PEG fragment interacted with the β-barrel through a direct hydrogen bond to E109, a water-mediated hydrogen bond to D122, and hydrophobic interactions with W190 (**Figure 6.10**). Whether these residues are also important for catalysis was investigated later. Like Pyrl4[159], Tsn15 also contains 14 amino acids at its *N*-terminus which could potentially act as a capping/lid-domain (as observed in Pyrl4[159]) (**Figure 6.11**).

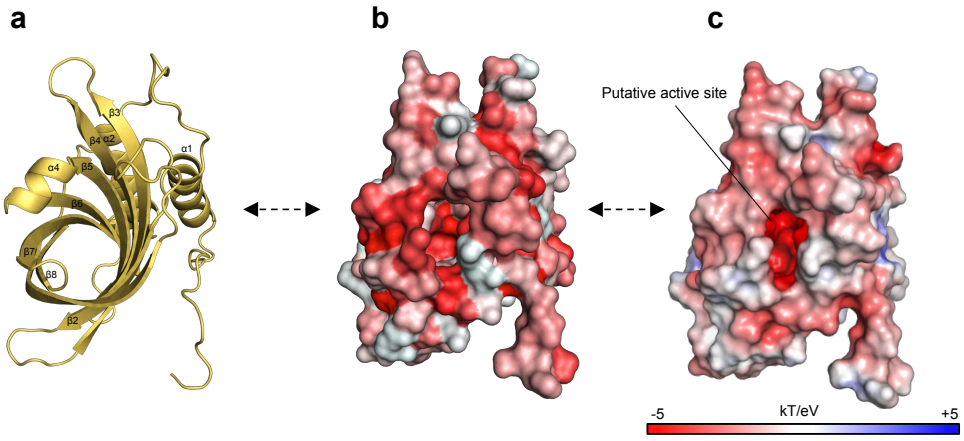


Figure 6.9 Hydrophobic cavity in Tsn15. a, Ribbon diagram of Tsn15, chain A. **b,** surface diagram of Tsn15, chain A with hydrophobic resides coloured red (darker red = more hydrophobic). c, Surface representation and surface electrostatics of Tsn15, chain A. Contouring at ±5 kT/eV. The putative active site cavity is indicated.

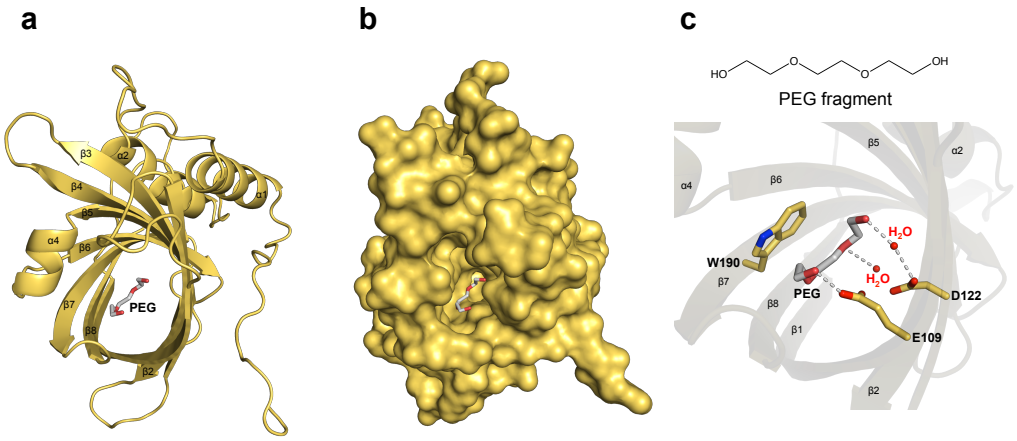


Figure 6.10 Bound PEG fragment in the β-barrel cavity of Tsn15. a, Ribbon representation of Tsn15 with a PEG fragment in its β-barrel cavity. **b,** Surface representation of Tsn15 with a PEG fragment in its β-barrel cavity. **c,** Amino acid residues involved in binding the PEG fragment. Grey dotted lines represent hydrogen bonds.

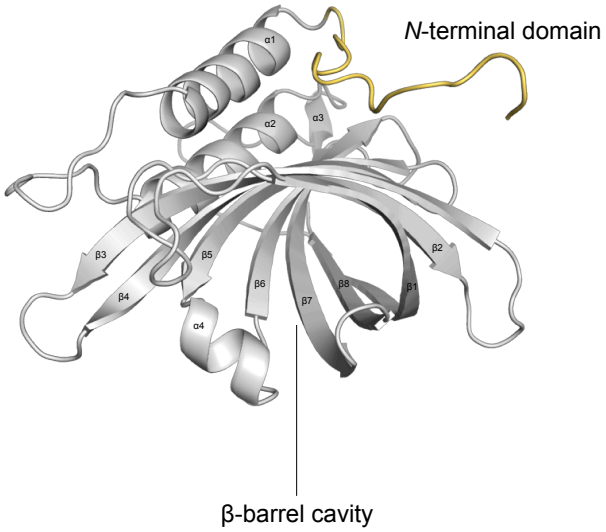


Figure 6.11 *N*-terminal region of Tsn15. The 14 amino acids of the *N*-terminal domain of Tsn15 are highlighted in yellow.

Tsn15 shares the same fold as its homologues AbyU and Pyrl4, and the allene oxide cyclases (such as AOC2) (**Figure 6.12**). Interestingly, Tsn15 was discovered to share its fold with PodA, a recently discovered pyocyanin demethylase from *Mycobacterium fortuitum*.

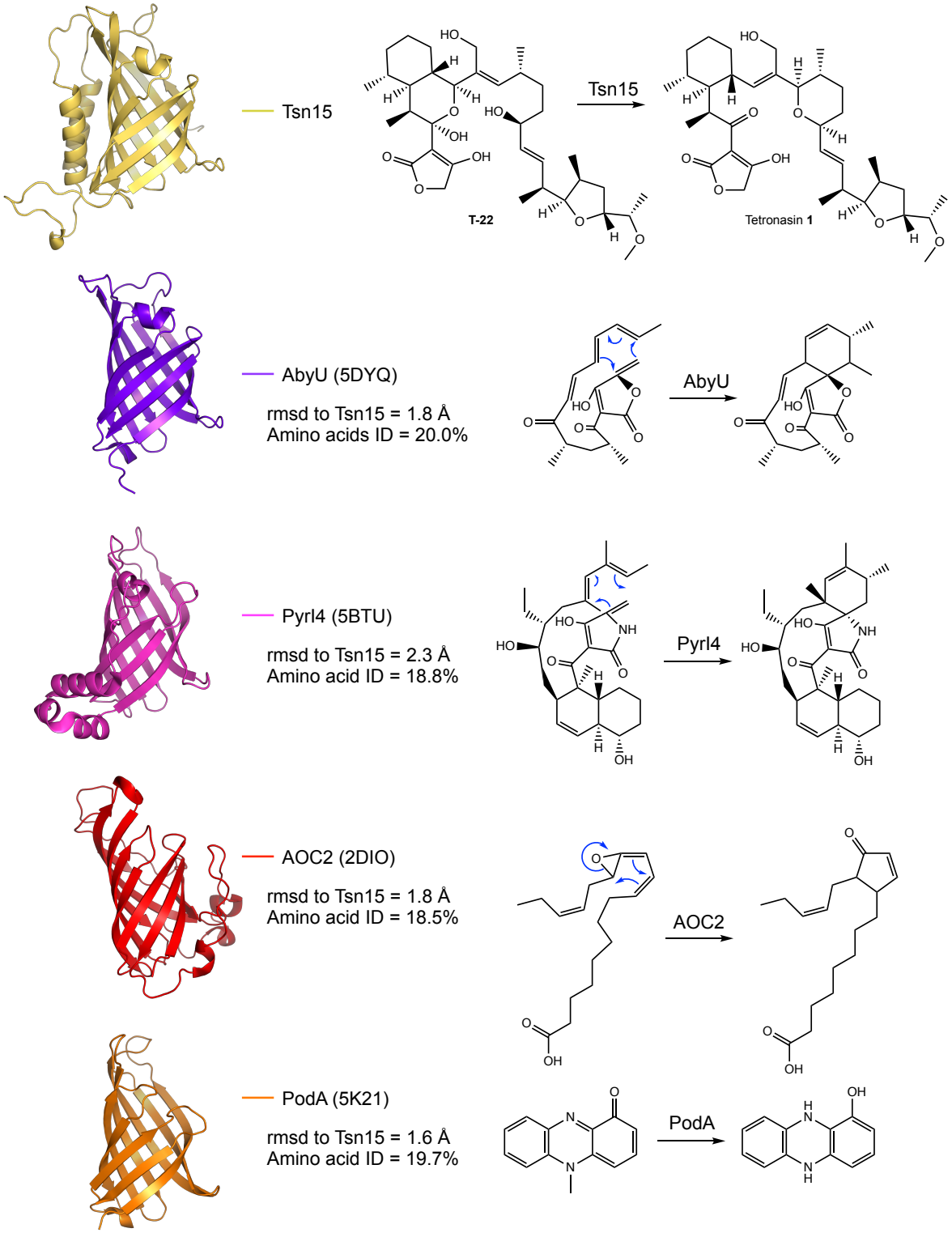


Figure 6.12 Structural homologues of Tsn15 and their respective reactions. The Cα chain of Tsn15 was aligned with its structural homologues AbyU, Pyrl4, AOC2, and PodA to calculate its rmsd value (alignment performed by Dr. Marcio Dias). Presented beside each structure is its rmsd value to Tsn15, amino acid identity (%) to Tsn15, and the reaction it catalyses.

6.4 Tsn15-ligand complex

While the structure of Tsn15 provided insight into its fold and dimerisation interface, it unfortunately did not provide any obvious mechanistic insights. The major reason being that how T-22 interacted with Tsn15 was unknown. It was therefore attempted to cocrystallise Tsn15 with T-22 in its active site. The success of such cocrystallisation experiments is unpredictable, as the substrate will often be converted into a product that does not bind with a high enough affinity to cocrystallise. T-22 purified by the author was sent to Fernanda Paiva, who performed the rest of the work described in this paragraph (figures were made by the author). Purified Tsn15 was mixed with T-22 before being added to a crystallisation drop (using the same crystallisation buffer as before). The Tsn15 crystals that grew were harvested and analysed at the PETRAIII synchrotron (Germany). From one of the crystals a 1.7 Å resolution dataset was collected. The phase problem was solved by molecular replacement using the previously solved Tsn15 structure. The structure of Tsn15 generated from the new crystal showed no significant differences from the first (**Appendix Figure 8**). However, there was additional electron density present in the β-barrel cavity of chain A, into which most of T-22 could be confidently modelled (mean B factor = 34 ± 11 $Å^2$) (collectively called the Tsn15-ligand complex) (**Figure 6.13**). The only part of T-22 missing electron density in the crystal was the C3 hemiacetal hydroxyl, indicating that T-22 had undergone a dehydration event either before binding or while bound to Tsn15. This dehydrated version of T-22 was labelled dT-22 and was also present in chain B, although in this chain the electron density from C17-C28 was too weak to accurately model (**Appendix Figure 9).**

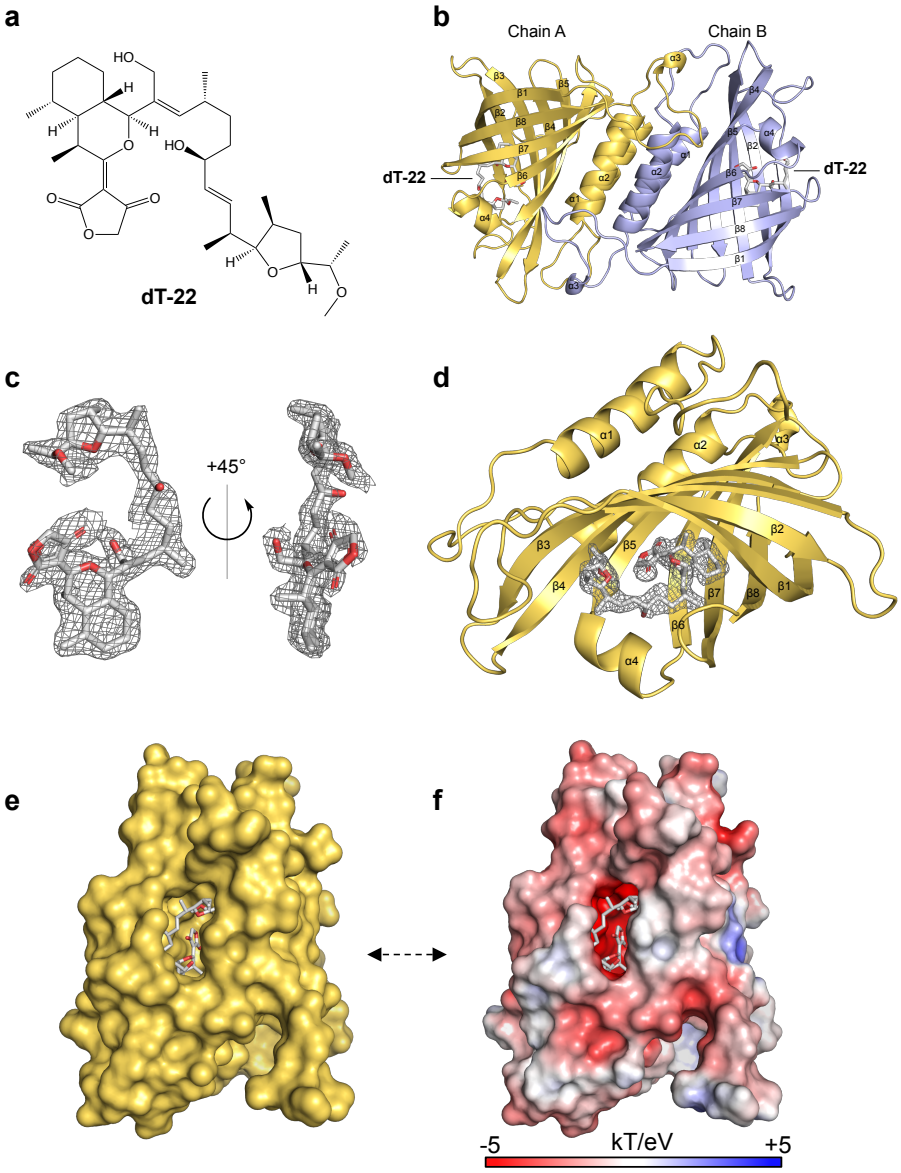


Figure 6.13 Structure of the Tsn15-ligand complex. a, The structure of dT-22. **b,** The structure of Tsn15-ligand complex was solved to 1.7 Å. Ribbon representation of the Tsn15-ligand structure with dT-22 bound in chain A. **c,** Three-dimensional structure of dT-22 overlaid with the calculated F_o−F_c electron density map (contoured at 2.0 σ). **d,** Ribbon diagram of Tsn15 containing dT-22 in its β-barrel cavity (F_o−F_c map contoured at 2.0 σ). **e,** Surface representation of Tsn15 containing dT-22 in its β-barrel cavity. **f,** Surface electrostatics of Tsn15 contoured at ±5 kT/eV.

The resolution of dT-22 was high enough to confirm that the stereochemistry of the chiral centres at C4, C5, C6, C10, C14, C17, C20, C21, C22, C24, and C25 are identical to tetronasin as predicted (**Section 5.2.4**). The tetronate ring of dT-22 is completely inserted into the β-barrel cavity of Tsn15, forming hydrogen bonds with S142 and Q164 (**Figure 6.14**). In chain B an additional hydrogen bonding interaction between R89 and the C30 hydroxyl is also present (**Appendix Figure 9**).

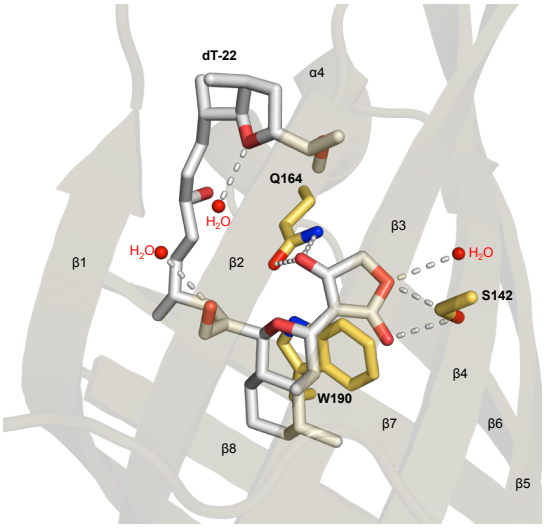


Figure 6.14 Polar interactions between dT-22 and Tsn15. Ribbon representation of the Tsn15-ligand complex (chain A). Hydrogen bonds are represented as dotted grey lines. Water molecules are presented as red spheres. The side chains of S142 and Q164 formed hydrogen bonds with the tetronate ring of dT-22.

The oxadecalin portion of dT-22 is oriented against the planar face of W190 through apparent π-stacking interactions (**Figure 6.14**). The tetrahydrofuran portion of dT-22 does not form any hydrogen bonds with amino acid side chains, likely making this portion of the molecule more flexible (consistent with its poor electron density in chain B). The conformation adopted by dT-22 does not appear to be the reactive conformation, as the atoms that must bond to form the tetrahydropyran ring, O17 and C13, are 4.6 Å apart (**Figure 6.15**). The reason for this is unknown, but numerous factors could be responsible, expanded upon in **Section 6.6**). Regardless, the apparent inert conformation of dT-22 bound to Tsn15 meant it was difficult to draw conclusions about the mechanism of Tsn15 without further experiments.

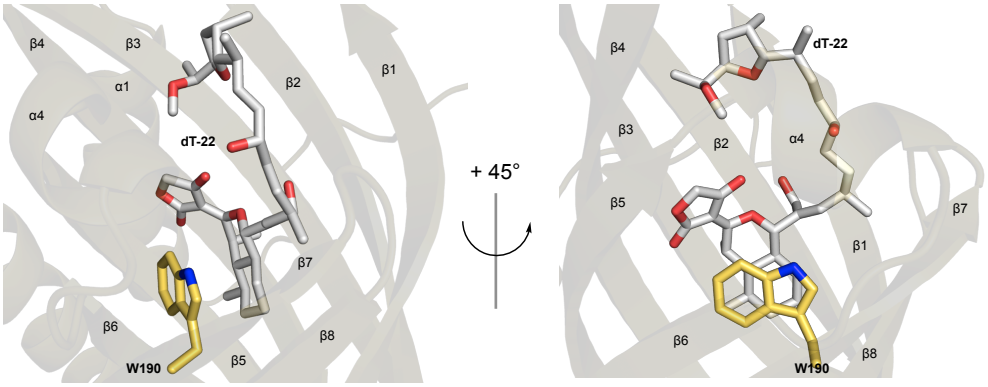


Figure 6.15 π-stacking interactions between W190 and the oxadecalin moiety of dT-22. Ribbon representation of Tsn15 chain A. The oxadecalin moiety of dT-22 is oriented parallel with W190 though apparent π-stacking interactions.

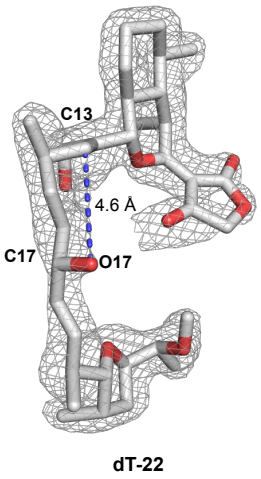


Figure 6.16 Distance between O17 and C13 in dT-22. In the conformation that dT-22 bound in the active site of Tsn15 O17 and C13 are 4.6 Å apart.

The proposed mechanisms for Tsn15 fell under two broad categories: first, Tsn15 may catalyse the nucleophilic attack of O17 onto C13 (mechanisms 5, 6, 7, and 8); alternatively, Tsn15 may catalyse a pericyclic rearrangement (mechanism 9 and 10). The nucleophilic attack mechanism was initially predicted as the most likely, as it would be similar to the mechanisms of other tetrahydropyran synthases such as Lsd19[300], MupZ[301], RED1[302], and SalBIII[197] (**Figure 6.16**). In all of these examples, a basic amino acid side chain extracts a proton from a hydroxyl group on the substrate, thereby increasing its nucleophilicity. An acidic side chain is often also present to act as a terminal proton donor.

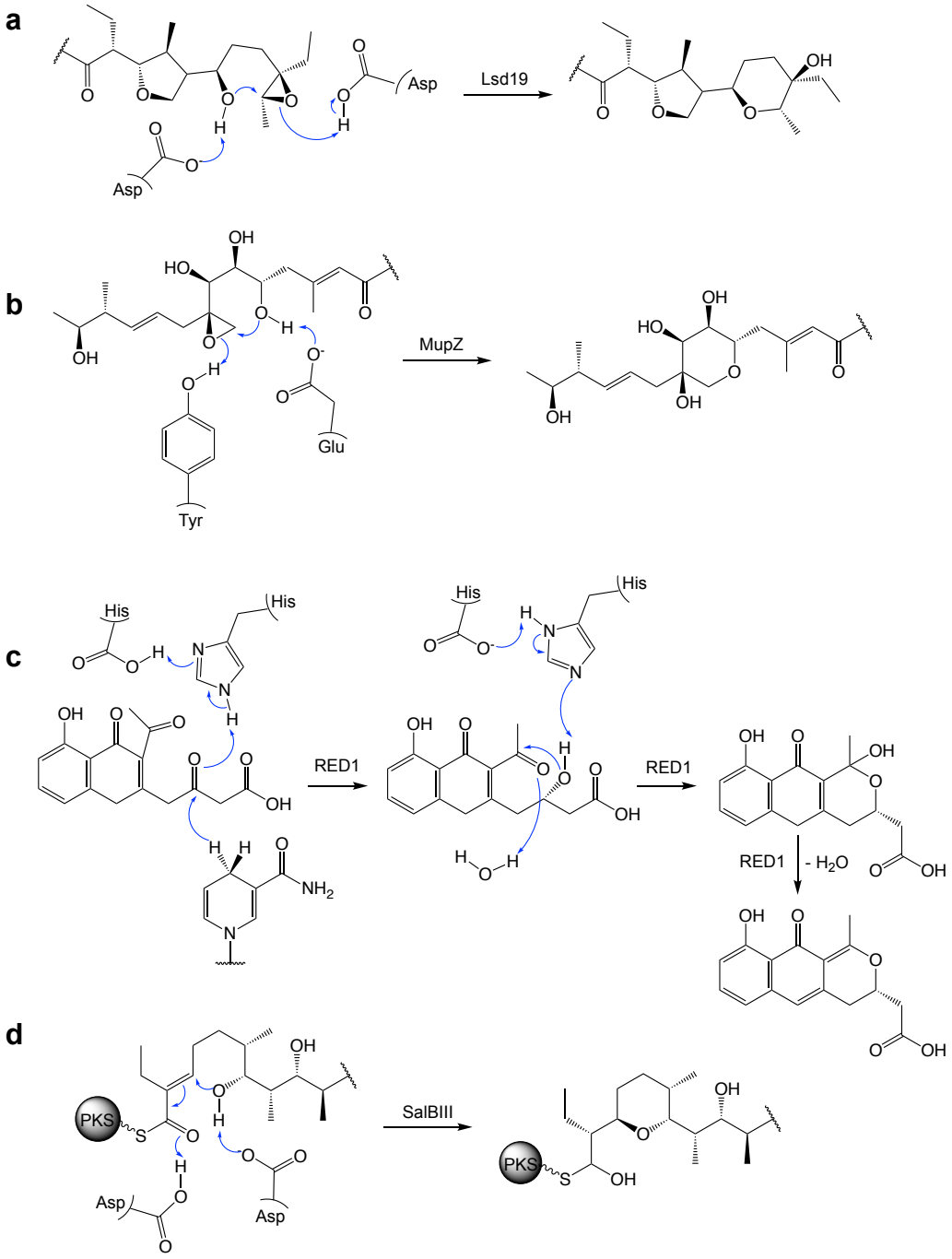


Figure 6.17 Enzymatic Mechanisms of tetrahydropyran formation. Proposed mechanism of: **a,** Lsd19 from lasalocid A biosynthesis[300]; **b,** MupZ from mupirocin biosynthesis[301]; **c,** RED1 in actinorhodin and dihydrogranaticin biosynthesis[302]; and **d,** SalBIII from salinomycin biosynthesis[197]. All of the enzymes presented use acid/base catalysis for tetrahydropyran formation.

As the active site of Tsn15 contains several acid/base residues (R89. E109, D122), it was hypothesised that these residues could be responsible for catalysing tetrahydropyran formation (**Figure 6.17**), even though in the Tsn15-ligand complex none of these acidic or basic side chains were adjacent to either O17 or the tetronate moiety of dT-22 (**Figure 6.18**).

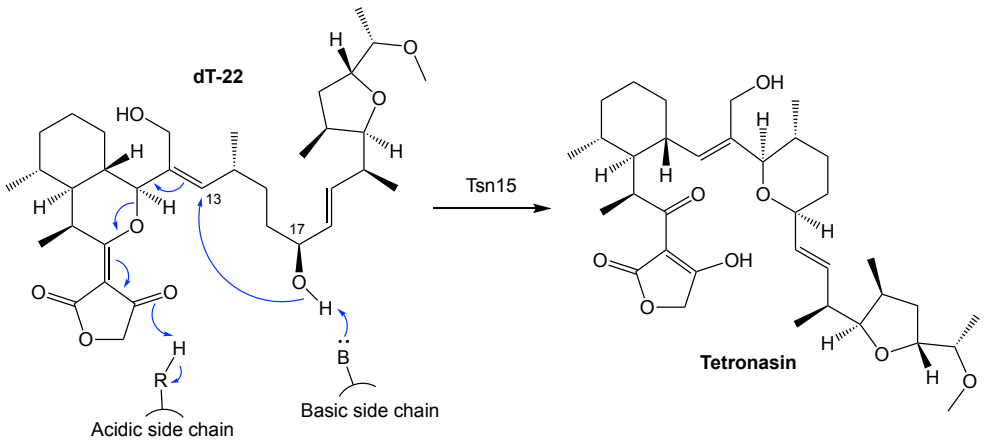


Figure 6.18 Potential Tsn15 acid/base-catalysed mechanism. A basic side chain that activates the O17 hydroxyl group could facilitate its attack onto C13. An acidic side chain adjacent to the tetronate ring could also assist in drawing electrons though the compound to close the tetrahydropyran ring and fragment the oxadecalin.

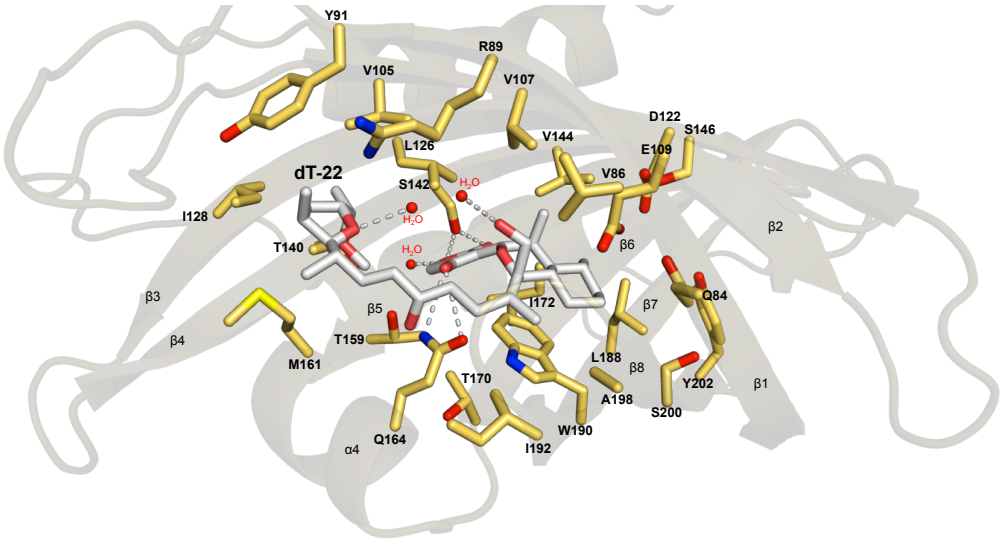


Figure 6.19 Amino acids present in the active site of Tsn15. Ribbon representation of Tsn15 (Chain A). The amino acids present on the internal face of the β-barrel are displayed as sticks. Hydrogen bonds are represented as dotted grey lines. Water molecules that form hydrogen bonds with dT-22 are represented as red spheres.

6.5 Site-directed mutagenesis of Tsn15

To probe the function of these and other residues site-directed alanine mutagenesis was performed (**Figure 6.19**). A total of 11 Tsn15 mutants were expressed and purified from *E.*

coli. In addition to acidic and basic amino acids, polar amino acids and other possible proton donors (such as tyrosine) were also mutated to alanine (or phenylalanine in the case of tyrosine). A truncated form of Tsn15 lacking *N*-terminal residues 1-11 (Tsn15-ΔN) was created to assess whether they form an essential substrate-docking lid domain, analogous to that proposed in PyrI4[159].

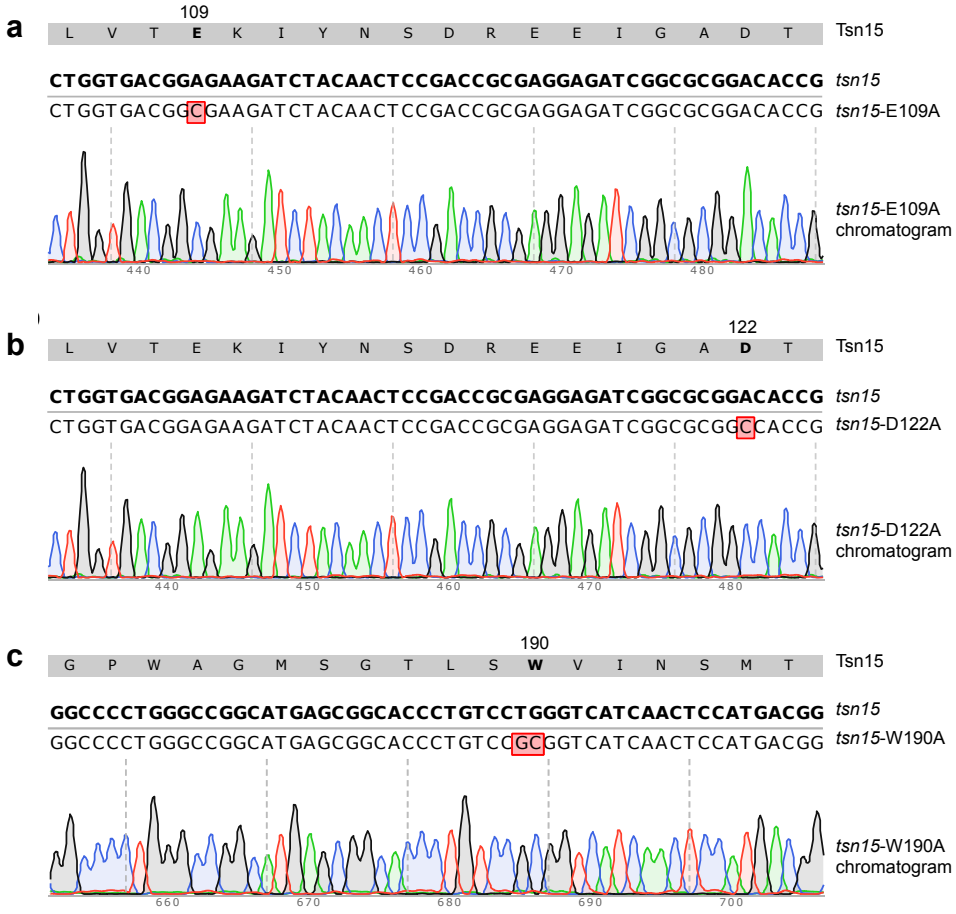


Figure 6.20 Creation of Tsn15 point mutants. a, DNA sequencing chromatograms confirming the successful creation of **a,** *tsn15*-E109A (GAG glutamic acid codon changed to a GCG alanine codon); **b,** *tsn15*-D122A (GAC aspartic acid codon changed to a GCC alanine codon); and **c,** *tsn15*-E109A (TGG tryptophan codon changed to a GCG alanine codon). Blue = cytosine; Red = thymine; Green = adenine; Black = guanine.

Each Tsn15 mutation was assayed by mixing it with T-16 and Tsn11 then measuring tetronasin production compared to a wild type Tsn15 control (**Figure 6.20**). These experiments demonstrated that none of the acidic/basic residues (R89A, E109A, D122A) nor Y91 or Y202 were essential for enzyme activity, suggesting Tsn15 does not use an acid/base mechanism. The pericyclic rearrangement (as depicted in mechanisms 9 and 10)

is therefore more likely (expanded upon at the end of the chapter). The Tsn15-ΔN mutant was still able to produce tetronasin. The only Tsn15 mutants displaying decreased activity compared to wild type Tsn15 were Q164A and W190. Tsn15 activity was completely abolished in the W190A mutant while Q164A mutant only had ca. 30% of the activity of Tsn15 wild type.

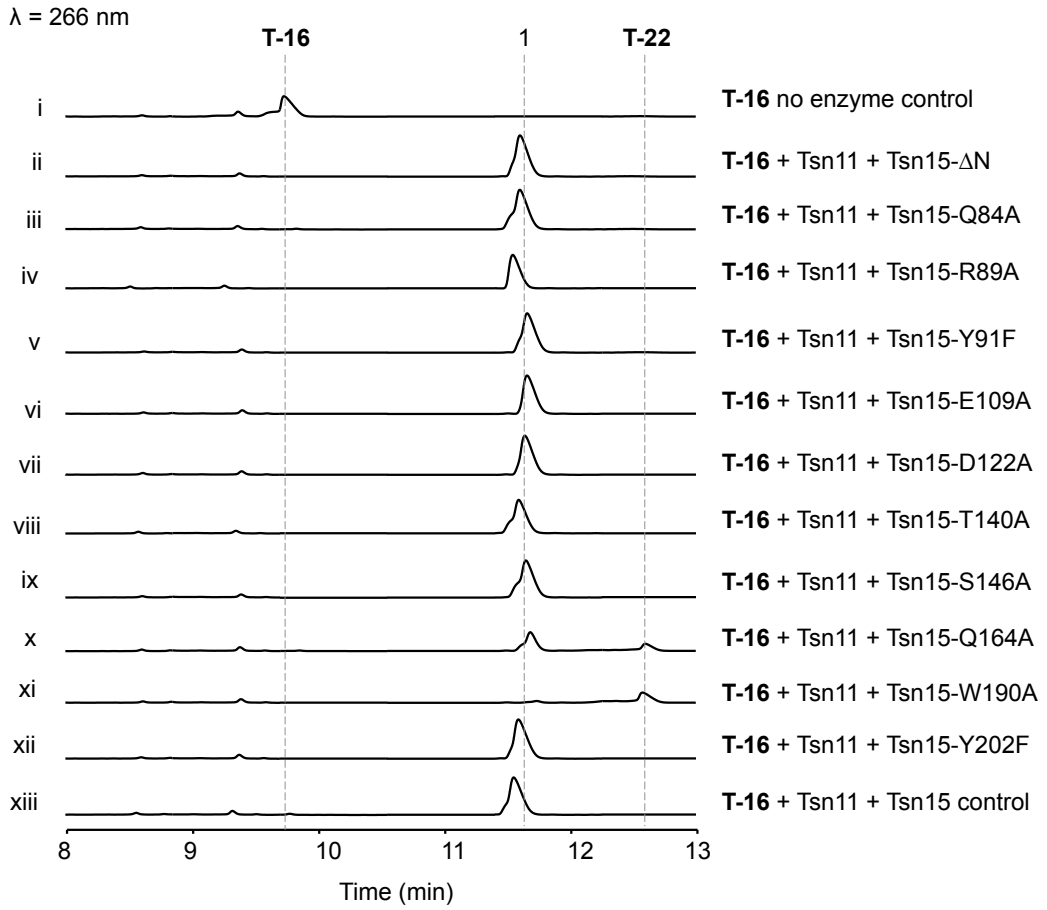


Figure 6.21 *in vitro* activity assays of Tsn15 mutants. T-16 was incubated with Tsn11 and a Tsn15 mutant. Production of T-22 and/or tetronasin (1) was monitored by HPLC (λ = 266 nm). (i) T-16 no enzyme control. Conversion of T-16 in the presence of Tsn11 and (ii) Tsn15 lacking its 11 *N*-terminal amino acids; (iii) Tsn15-Q84A; (iv) Tsn15-R89A; (v) Tsn15-Y91F; (vi) Tsn15-E109A; (vii) Tsn15-D122A; (viii) Tsn15-T140A; (ix) Tsn15-S126A; (x) Tsn15- N151A; (xi) Tsn15-Q164A; (xii) Tsn15- W190A; (xiii) Tsn15- Y202F; (xiv) Tsn15 wild type control. Data are representative of three independent experiments.

The decreased activity of these Tsn15 mutants is likely due to disrupted substrate binding; W190 normally making π-stacking interactions and Q164 normally hydrogen bonding (**Figure 6.14**). A tryptophan residue is present at the equivalent locus in both Tmn8 and Mad31 (**Figure 6.21**), suggesting a conserved role in binding to the oxadecalin intermediates of tetronomycin, and tetromadurin, respectively (**Figure 6.21**). Unfortunately, a S142A mutant could not be created in the time frame of this work; however, this residue

is likely also important for substrate binding as it hydrogen bonds with the tetronate ring. An equivalent serine (or threonine) is conserved in both Tmn8 and Mad31, indicating that these residues may also have a conserved role in bonding to the tetronate ring of the equivalent tetronomycin and tetromadurin intermediates.

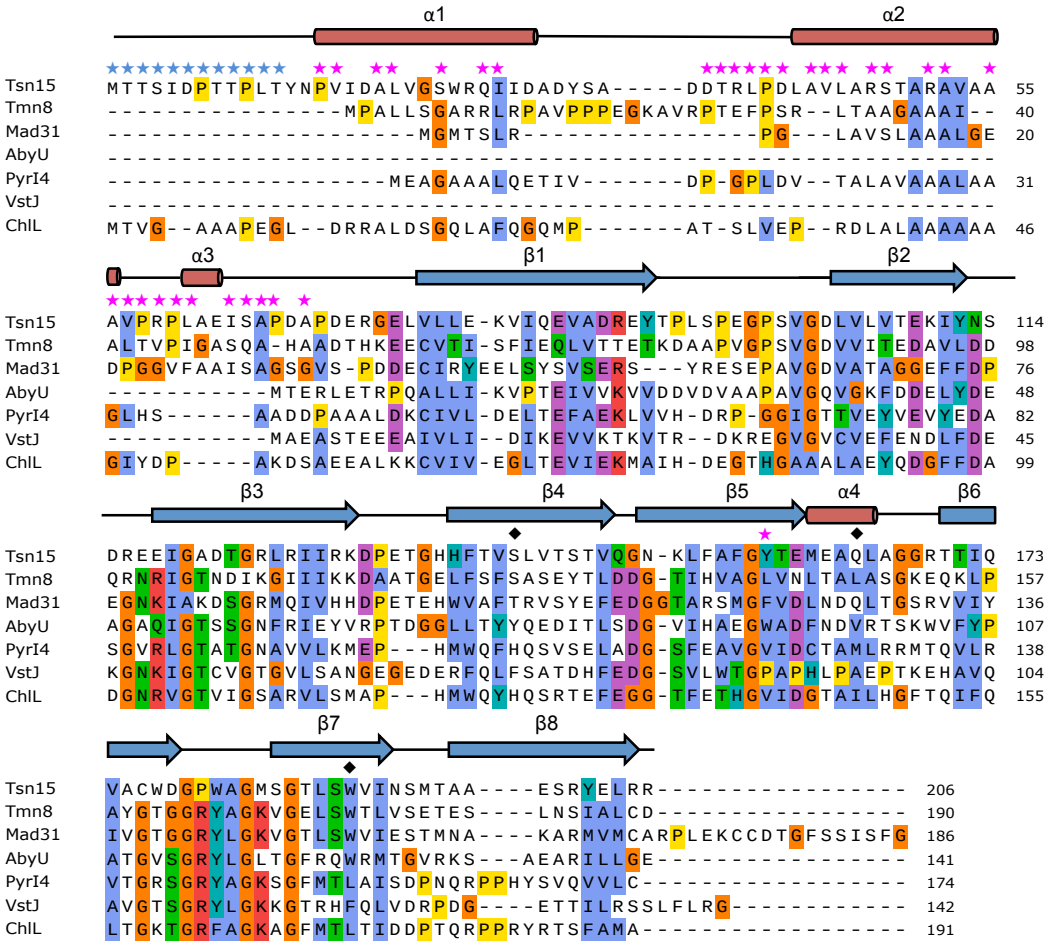


Figure 6.22 Secondary structure and multiple sequence alignment of Tsn15. Tsn15 was aligned with its homologues from polyether tetronate (Tmn8 and Mad31) and spirotetronate (AbyU, PyrI4, VstJ, and ChlL) pathways using ClustalOmega[227]. The secondary structure of Tsn15 is presented above its sequence. The *N*-terminal residues removed in the Tsn15-ΔN mutant are highlighted with blue stars (★). Amino acid residues involved in the dimerisation interface are highlighted with magenta stars (★). Amino acids involved in binding the dT-22 complex are highlighted with back diamonds (◆).

To check that the decreased activity of the W190 and Q164 mutants was due to the amino acid substitution, rather than a major change in Tsn15 conformation, circular dichroism (CD) spectra were collected. Circular dichroism is a spectroscopic method for analysing protein secondary structure. In CD spectroscopy a beam of light is polarised into clockwise and counter clockwise rotating waves[303]. Asymmetric molecules such as proteins absorb

the two circularly polarised light waves to different extents. Crucially, α-helices and β-sheets each have characteristic CD absorption spectra[303], enabling their detection in a protein of interest. Comparing the CD spectra of Tsn15-Q164 and Tsn15-W190 revealed they had no significant differences in protein secondary structure compared to wild type Tsn15 (**Figure 6.22**). The decreased activity of these mutants is therefore due to the amino acid substitution alone.

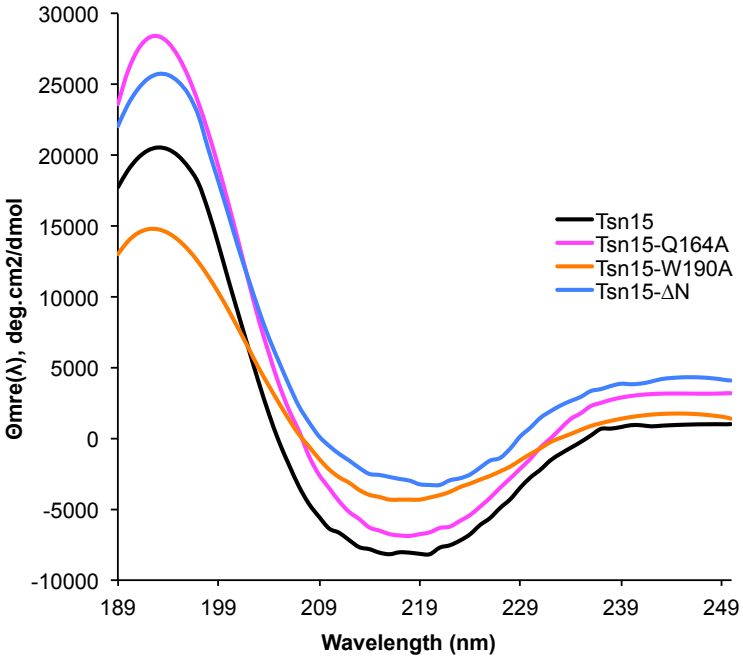


Figure 6.23 Circular dichroism (CD) spectra of Tsn15 mutants. CD spectra of Tsn15 (black), Tsn15-Q164A (magenta), Tsn15-W190A (orange), and Tsn15-ΔN.

6.6 Chapter 6 Discussion

This work describes the first structural and mechanistic studies on the [4+2] cyclase homologues from a polyether tetronate biosynthetic pathway. Like many of their homologues, both Tsn11 and Tsn15 were determined to be homodimeric in solution (**Figure 6.1**)[158–160,163,259,293]. Based on its similarity to PyrE3, each monomer of Tsn11 likely binds a single molecule of FAD[163]. Both proteins were crystallised in commercially available, premade, crystallisation screening plates (**Figure 6.4**). The FAD-binding Tsn11 crystallised as long, thin, yellow protein crystals that diffracted poorly (>4 Å), meaning its structure could not be solved. The only example of an $_{INV}$HDA-catalysing enzyme with a solved protein structure solved is LepI from leporin biosynthesis[280,304], though its

mechanism is still unclear[305]. Solving the structure and understanding the mechanism of Tsn11 is therefore of great interest. Furthermore, if the mechanism of LepI is eventually discovered it could lead to interesting comparisons with Tsn11, providing insights into the convergent evolution of these two enzymes. Several strategies could assist in obtaining better diffracting Tsn11 crystals in the future. First of all, the ÄKTA trace of Tsn11 indicates that it is not entirely monodisperse in solution (**Figure 6.3**). The low levels of non-homodimeric Tsn11 could have interfered with protein packing during the crystallisation process, resulting in a poorly diffracting crystal. A future experiment could test a range of different protein buffers to identify one in which Tsn11 remains in a single oligomeric state better. Additionally, truncated forms of Tsn11 could be crystallised and screened for their diffraction potential. The crystal structure of PyrE3 revealed that its *C*-terminal region forms a thioredoxin-like domain apparently not involved in catalysis[163]. From the alignment of Tsn11 with PyrE3, it is likely that Tsn11 shares this same domain (**Figure 3.21**). Removing the thioredoxin-domain of Tsn11 could therefore be performed in an attempt to reduce the flexibility of the protein and improve crystal packing.

Unlike Tsn11, the structure of Tsn15 could be solved to 1.8 Å using SAD phasing. Following this, Fernanda Paiva cocrystallised Tsn15 with T-22 to solve the structure of a Tsn15 with dT-22 bound in its active site. The amino acids S142, Q164, and W190 were identified as being key substrate binding residues (**Figure 6.13, 6.20**). That Tsn15 contains a specific residue (W190) for interacting with the oxadecalin moiety of T-22 provides additional evidence against mechanism 8 (**Figure 5.31**) (where the minor, open ring, component of T-22 is the Tsn15 substrate) being the true mechanism. Fragmentation of the oxadecalin to form the cyclohexane ring of tetronasin could decrease the strength of the W190 π-stacking interactions, preventing product inhibition. That the tetrahydrofuran portion of T-22 was facing out of the β-barrel cavity towards the solvent could explain why changes in this part of the molecule do not affect Tsn15 catalysis. As discussed in Chapter 5, the cyclohexane and tetrahydropyran rings of tetronasin can still form if the tetrahydrofuran ring is modified (or in an C24-C25 *E* alkene open form).

Tsn15 shares its fold with enzymes possessing a diverse range of activities, including the [4+2] cyclases PyrI4[159] and AbyU[158], the allene oxide cyclases[160,296], and the demethylase PodA[293]. These different homologues highlight the versatility of this enzyme fold for catalysing a diverse range of reactions, including non-pericyclic ones. Research into the

evolution of these enzymes/conserved folds would be an interesting future topic, particularly if more examples of the fold are characterised. The versatility of this enzyme fold could also make it a useful scaffold for developing novel catalysts for synthetic biology with a diverse range of possible activities[158].

The Tsn15 mutagenesis experiments established unequivocally that certain residues that were candidates to supply acid-based catalysis are not essential. In contrast, in the Tsn15 W190A, mutant activity is completely abolished. The Q164A mutant displayed a 30% decrease in activity, but whether this reflects a change in enzyme affinity or rate can only be speculated. The tolerance of Tsn15 to mutations would assist in engineering the enzyme to accept substrate analogues to produce novel polyether tetronates. Unfortunately, due to time constraints and the limited amounts of the T-16 and T-22 substrates available, K_m and k_{cat} values for Tsn15 and Tsn11 were not measured, and this ought to be addressed in future work on those mutants where activity is altered.

That dT-22 crystallised in the active site of Tsn15, rather than T-22 or tetronasin, complicated the mechanistic analysis. Further confounding was that dT-22 appears to be in an inert conformation in the Tsn15 active site. However, if dT-22 is the substrate of Tsn15, then its inert conformation in the crystal must explain its persistence there. Another possibility is that T-22 is the true substrate of Tsn15 and dT-22 is an inert spontaneous dehydration product (**Figure 6.23b**). While dT-22 not being a substrate of Tsn15 could explain its presence in the Tsn15 crystal, the direct conversion of T-22 into tetronasin has problems of its own. For T-22 to be converted into tetronasin in a single enzymatic step, the hemiacetal hydroxyl must act as a leaving group. As hydroxyl groups are poor leaving groups, it would almost certainly need to accept a proton for this mechanism to be possible (as water is a much better leaving group) (**Figure 6.23a**).

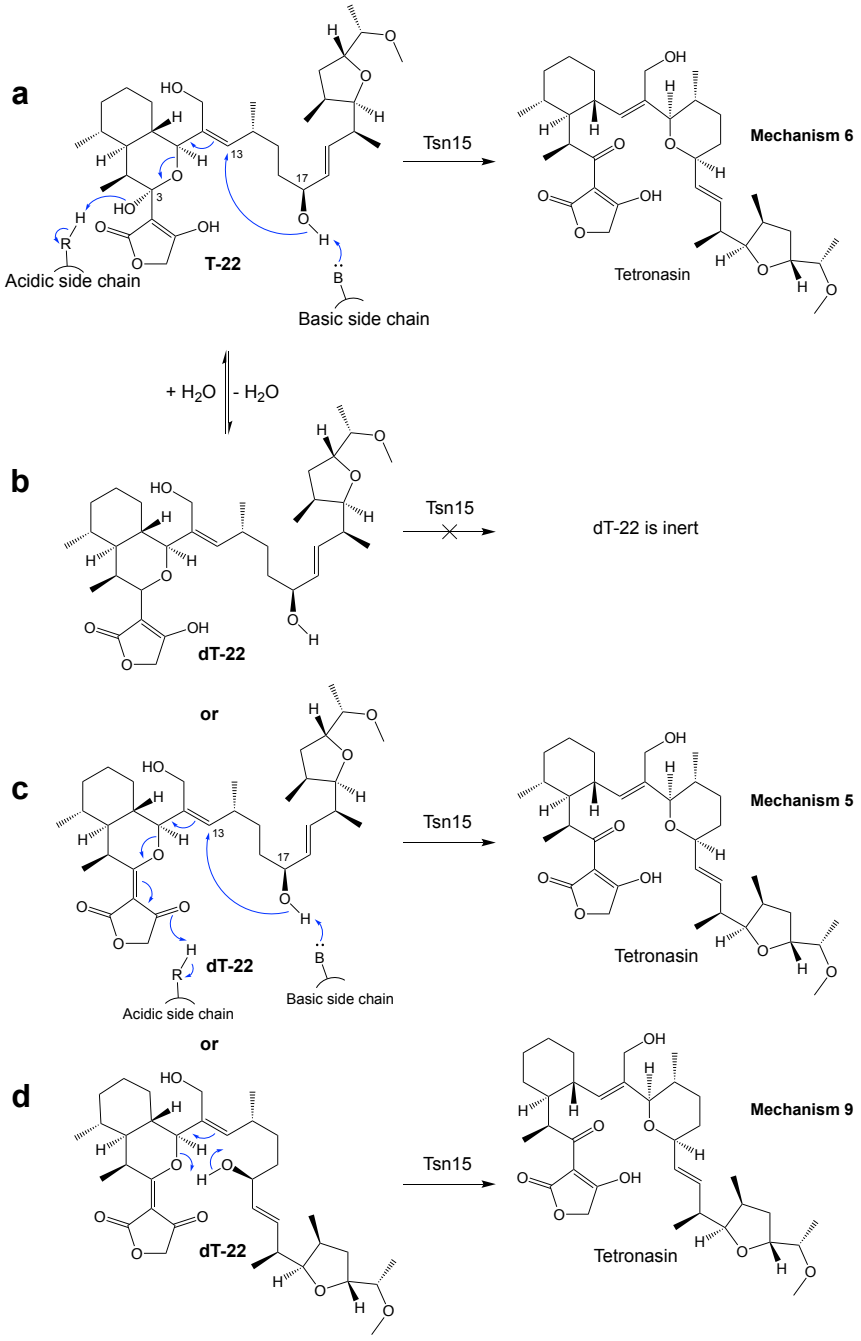


Figure 6.24 Possible mechanisms for Tsn15-catalysed formation of the tetronasin tetrahydropyran ring. a, In mechanism 6, T-22 is the direct substrate of Tsn15 and tetrahydropyran formation and hemiacetal dehydration occur simultaneously. An acidic group would likely be required to protonate the hemiacetal hydroxyl to make it a better leaving group. A basic side chain could also help the reaction by activating the C17 hydroxyl nucleophile. **b,** Another possibility is that dT-22 could be an inert spontaneous dehydration product of T-22, explaining why it co-crystallised in the active site of Tsn15. **c,** In mechanism 5, dT-22, rather than T-22, is the substrate of Tsn15. It is likely that this mechanism would also require a basic side chain to activate the C17 hydroxyl nucleophile **d,** In mechanism 9, Tsn15 catalyses a pericyclic reaction to form the tetrahydropyran ring of tetronasin. Acidic and basic amino acid side chains are not necessarily required for this mechanism to occur.

The source of this proton would likely be an acidic amino acid side chain in the active site. A basic side chain (such as a histidine) could also accelerate this reaction by extracting a proton from the C17 hydroxyl, making the oxygen a stronger nucleophile. However, the mutagenesis experiments demonstrated that the potential proton-donating side chains present in the active site (R89, Y91, E109, and D122, and Y202) are not essential for Tsn15 activity (**Figure 6.20**). These data argue against the notion that T-22 is the substrate of Tsn15. However, if T-22 is not the substrate of Tsn15, then it suggests that dT-22 is. The dehydration of T-22 to dT-22 could occur spontaneously, rather than catalysed by Tsn15, through a sterically favourable six-membered transition state (**Figure 6.24**). Whether Tsn15 accelerates this dehydration, and its reversibility, is unknown.

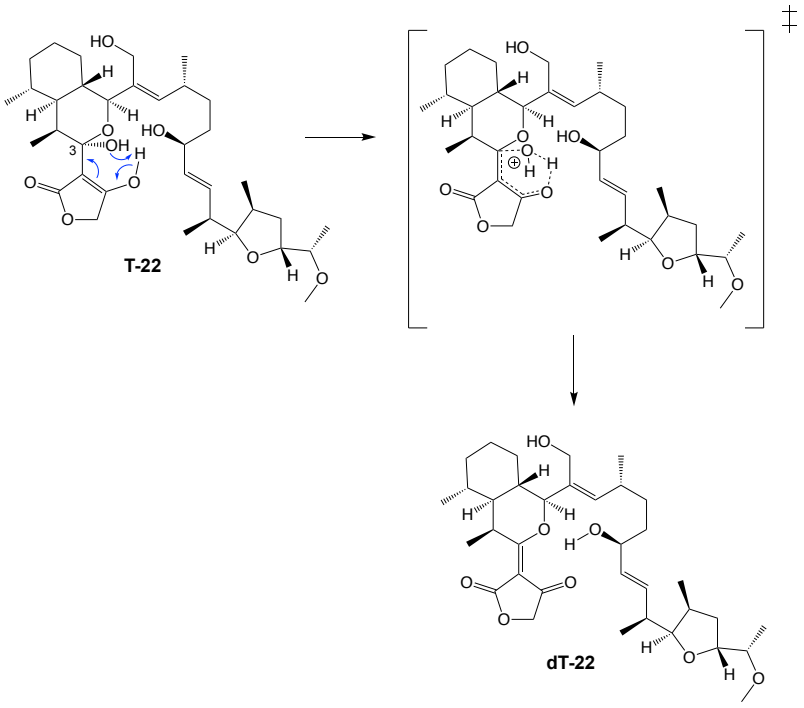


Figure 6.25 Proposed mechanism of spontaneous dehydration of T-22 to form dT-22. The dehydration of the C3 hemiacetal hydroxyl could occur via a six-membered transition state.

However, if dT-22 is the true substrate of Tsn15, there is still the question as to why it appears to be in an inert conformation (**Figure 6.15**). There are several possible explanations for this, one being that the packing of the crystal lattice imposed conformation restraints on Tsn15, preventing dT-22 from adopting its reactive conformation. Alternatively/in addition, the overabundance of T-22 compared to Tsn15 used in the cocrystallisation process (10 mM of T-22 to 0.3 mM of Tsn15) likely meant the equilibrium

favoured the substrate, even if conversion to tetronasin was also occurring. As the substrate should have a higher binding affinity to Tsn15 than the product, this could have led to dT-22 co-crystallising in the active site.

Assuming dT-22 is the true substrate of Tsn15, there is still the question as to which of the proposed mechanisms is the most likely. The Tsn15 mutagenesis experiments indicated that a general acid/base-catalysed mechanism, as presented in mechanisms 5, 6, 7, and 8 is unlikely (Figure 5.31, **Figure 6.23c**). In the remaining mechanisms, 9 and 10, Tsn15 catalyses a pericyclic rearrangement to form the tetrahydropyran ring and fragment the decalin ring, forming tetronasin. The pericyclic rearrangement could either occur after hemiacetal dehydration (mechanism 10), or before (mechanism 9). As established, Mechanism 10 is unlikely, as it would result in an oxygen atom derived from water being incorporated into C3 keto of tetronasin. Previous isotopic feeding experiments have shown that this does not happen[175]. From the data collected in this thesis, mechanism 9 (redrawn in figure 6.25) therefore appears the most likely. Mechanism 9 has an additional advantage as homologues of Tsn15 are known to catalyse pericyclic reactions in other pathways (such AbyU[158], PyrI4[159], and AOC2[160,296]). A pericyclic mechanism would also not necessarily depend on acidic/basic side chains, only that the substrate be oriented in a correct conformation for the reaction to occur. The inability to identity essential acidic/basic amino acids in the active site of Tsn15 therefore provides additional support (**Figure 6.20**). Essential amino acid residues were also not identified in active sites of the Tsn15 homologues AbyU[158] and PyrI4[159], and the unrelated [4+2] cyclases PyrE3[163] and SpnF[155]. While the mechanism appears the most likely (**Figure 6.25**), the alternatives cannot be ruled out without additional experiments. A valuable experiment would be obtaining evidence for the formation of dT-22 during the Tsn15 reaction with T-22, or capturing dT-22 directly and testing its competence as a Tsn15 substrate. Either of these experiments could rule out T-22 being the direct substrate of Tsn15.

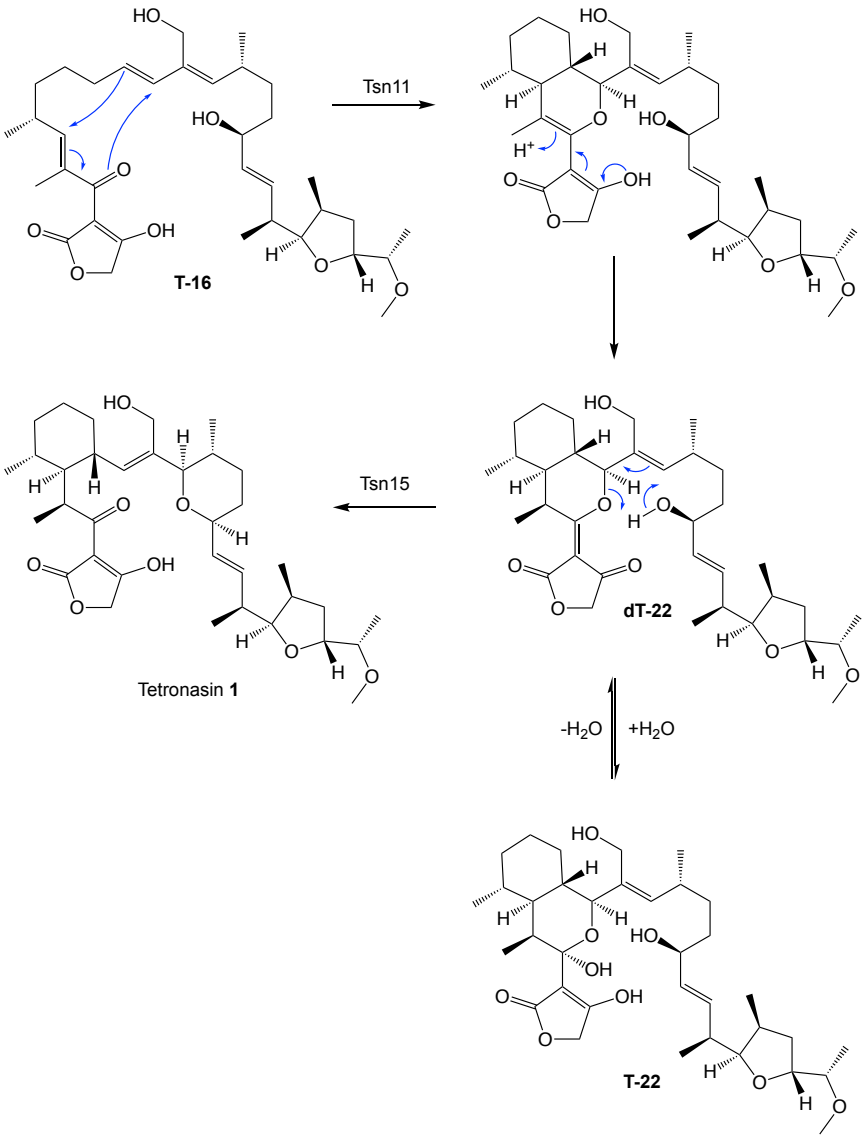


Figure 6.26 Proposed mechanism for formation of the cyclohexane and tetrahydropyran rings of tetronasin. Tsn11 catalyses an $_{INV}$HDA reaction to form an oxadecalin intermediate. The oxadecalin is then hydrated to form the cyclic hemiacetal T-22. T-22 then undergoes a dehydration reaction to form dT-22 (it is unknown whether this occurs spontaneously or is catalysed by Tsn15), which is converted into tetronasin **1** via a pericyclic reaction catalysed by Tsn15.

Chapter 7: Key findings and future directions

7.1 Research motivation

The primary goal of this research was to understand how the cyclohexane and tetrahydropyran rings are formed in the polyether tetronate ionophores tetronasin, tetronomycin, and tetromadurin. Although research on polyether tetronate biosynthesis, tetronasin in particular, began over 20 years ago, how these two rings form remained a mystery. Different theories were proposed, the most prominent being that the cyclohexane and tetrahydropyran rings are formed in a single step via metal catalysis[175,177]. In 2015, a study on the biosynthesis of the spirotetronate versipelostatin provided the first hint that an enzyme is instead responsible[157]. A single enzyme, VstJ, was found to catalyse formation of the spirotetronate moiety of versipelostatin via an apparent [4+2] cycloaddition (Diels-Alder reaction)[157]. Unexpectedly, a homologue of *vstJ* was also identified in the biosynthetic gene cluster of tetronomycin, one of the polyether tetronates[111,157]. The function of *tmn8* was unknown, as none of the polyether tetronates possess a spirotetronate moiety. An exciting possibility was that Tmn8, although resembling a [4+2] cyclase, was actually responsible for catalysing cyclohexane and/or tetrahydropyran formation. Polyether natural products, in addition to often being antibiotics, have recently demonstrated potential as anticancer drugs[173]. Discovering the enzyme(s) responsible for the cyclohexane and tetrahydropyran formation in the polyether tetronates could enable them to be used as synthetic biology tools for synthesising novel analogues. As such, there are clear incentives for understanding the intricacies of polyether tetronate biosynthesis.

7.2 Key findings

The first goal of this thesis was to determine if *vstJ* homologues are also present in the BGCs of tetronasin and tetromadurin. To achieve this, the tetromadurin BGC (the *mad* cluster) in the genome of *A. verrucosispora* was first identified (**Chapter 3**). The *mad* cluster was immediately identifiable, containing the necessary type I PKS genes to produce the tetromadurin carbon backbone and accessory genes for the biosynthesis of its key features. Like tetronomycin, tetronate ring biosynthesis for tetromadurin uses a glyceryl-ACP unit. Tetronasin therefore remains the only known polyether tetronate (and tetronate natural product in general) to use glycolyl-ACP for tetronate biosynthesis. The carbon backbone of tetromadurin is produced by seven type I PKS (MadAI-MadAVII), together contributing a loading module and 14 extension modules that correspond well to the theoretical linear intermediate of tetromadurin. The one exception was module 12, which contains an apparent inactive KR domain despite this module apparently reducing the C7 β-keto group of the polyketide intermediate. How reduction of the C7 keto group occurs is still unclear, but could involve a domain from an adjacent module.

Next, BLAST was used to successfully identify *vstJ* homologues in the *mad* and *tsn* clusters. In addition, genes homologous to a second class [4+2] were also identified in the BGC of all three polyether tetronates. These genes were homologues of *pyrE3*, a recently characterised [4+2] cyclase responsible for dialkyldecalin formation in pyrroindomycin biosynthesis[161]. Intriguingly, the gene encoding the PyrE3 homologue in the *tmn* cluster, *tmn9,* had previously been shown to be essential for tetronomycin biosynthesis, albeit for unknown reasons[111].

Coding-frame gene deletions of each [4+2] cyclase homologue were performed to investigate their role in the polyether tetronate biosynthesis (**Chapter 4**). Deletion of the *vstJ* homologues *tmn8*, *tsn15,* and *mad31* abolished biosynthesis of their respective polyether tetronate almost entirely. Likewise deletion of the *pyrE3* homologues *mad10* and *tsn11* also abolished tetromadurin and tetronasin production, respectively. Both classes of [4+2] cyclase homologue are therefore required for polyether tetronate biosynthesis. The *S. longisporoflavus* Δtsn11 mutant produced a new metabolite (T-16) with the same molecular weight as tetronasin, consistent with it being a precyclisation form of tetronasin lacking the cyclohexane and tetrahydropyran rings. T-16 was also produced by the *S. longisporoflavus*

Δtsn11 Δtsn15 double mutant, indicating that Tsn15 can only act after Tsn11. The next step was to enough isolate T-16 for structural characterisation by NMR.

Approximately 12.5 mg of T-16 was isolated from an upscaled fermentation of *S. longisporoflavus* Δtsn11 (**Chapter 5**). A purification protocol was developed using C18 solid phase extraction cartridges followed by semipreparative HPLC. Detailed NMR characterisation of T-16 confirmed it as a tetronasin analogue in which the tetronate and tetrahydrofuran rings had formed, but the cyclohexane and tetrahydropyran rings had not. When T-16 was incubated with Tsn11 it was converted into a new metabolite, labelled T-22. T-22 could be subsequently converted into tetronasin by incubating it with Tsn15. Tsn11 and Tsn15 therefore act in an enzymatic cascade to form the cyclohexane and tetrahydropyran rings of tetronasin. The Tsn11 *in vitro* assay was scaled up to isolate ca. 3 mg of T-22 for NMR structural characterisation. T-22 still lacked the tetrahydropyran ring of tetronasin, but the cyclohexane ring had closed within a larger oxadecalin moiety. The oxadecalin moiety was unexpected but is consistent with Tsn11 catalysing an apparent $_{INV}$HDA reaction, reminiscent of the reaction catalysed by PyrE3[161]. Notably, the product of this apparent [4+2] cycloaddition, the oxadecalin moiety, is not present in the final tetronasin structure, meaning Tsn11 catalyses a latent reaction. Tsn15 then catalyses tetrahydropyran ring formation and fragments the oxadecalin moiety to form tetronasin. Tsn11 and Tsn15 therefore catalyse different reactions from their homologues in spirotetronate/spirotetramate pathways. The same enzymatic cascade mechanism is likely responsible for forming the cyclohexane and tetrahydropyran rings of tetronomycin and tetromadurin. Collaborators from the University of Warwick used chemical probes to offload tetronasin intermediates prematurely, indicating that the tetrahydrofuran ring likely forms prior to chain release. The final proposed order of ring formation is therefore 1) tetrahydrofuran, 2) tetronate, 3) cyclohexane, 4) tetrahydropyran.

Despite the structure similarities between tetronasin, tetronomycin, and tetromadurin, Tmn8 and Mad31 could not substitute for Tsn15 *in vivo* or *in vitro*. Likewise Tmn9 was unable to substitute for Tsn11 (Mad10 could not be purified, unfortunately). Like PyrE3, Tsn11 resembles an FAD-dependent monooxygenase and binds an FAD cofactor. Chemically reducing this FAD cofactor using sodium dithionite did not affect Tsn11 activity, indicated that it does not have a catalytic role.

Finally, Tsn15 was crystallised and its structure was solved at 1.8 Å resolution (**Chapter 6**). Tsn15 is a homodimer with each monomer comprised of an eight-strand β-barrel and a α-helix based dimerisation domain. Despite a low sequence identity (ca. 20%) Tsn15 shares the same fold as the [4+2] cyclases PyrI4[159] and AbyU[158], the allene oxide cyclases[160,296], and, noted for this first time in this work, the pyocyanin demethylase PodA[293]. In attempt to gain insight into the mechanism of Tsn15, collaborator Fernanda Paiva (University of São Paulo, Brazil) cocrystallised T-22 with Tsn15. Within the β-barrel cavity of Tsn15 a dehydrated form of T-22 (dT-22) had cocrystallised, enabling key substrate binding residues (S142, Q164, and W190) to be identified. S142 and Q164 both formed hydrogen bonds to the tetronate ring, while W190 made π-stacking interaction with the oxadecalin portion of dT-22. Despite the insight gained into substrate binding, the mechanism of Tsn15 remains uncertain. dT-22 appears bound to Tsn15 in an inert conformation, so it is possible that T-22 is the real substrate and dT-22 is a spontaneous dehydration shunt product. Regardless, there are two main possible mechanisms for Tsn15: acid/base catalysis or pericyclic ring closure. The acid/base catalytic mechanism appears unlikely as acidic/basic amino acids in the active site were shown by mutagenesis to not be essential for Tsn15 activity. The alternative is that Tsn15 closes the tetrahydropyran ring/opens the oxadecalin ring via a pericyclic rearrangement. The lack of essential acid/base side chains in the active site therefore favours the pericyclic mechanism, though additional experiments are necessary to confirm.

7.3 Future Directions

An area of this thesis that clearly requires expansion upon is the proposed mechanisms of Tsn11 and Tsn15. Continued efforts to crystallise Tsn11 to solve its structure would be valuable, as a crystal structure would be the best starting point for mechanistic analysis. Even better would be if Tsn11 were cocrystallised with its substrate, T-16. If the cocrystallisation was unsuccessful an alternative could be to try docking T-16 computationally. Site-directed mutagenesis could then be performed to probe the putative binding pocket of Tsn11 for catalytic residues. A docking approach was used for the Tsn11 homologue, PyrE3 after cocrystallisation with its substrate failed[163].

Additional experiments are also required to elucidate the mechanism of Tsn15. An important milestone in future work will be to determine whether T-22 or dT-22 is its true substrate. To study the mechanism in more detail Fernanda Paiva is planning to conduct comprehensive molecular dynamic (MD) simulations of the Tsn15-ligand complex. These simulations will predict the conformational flexibility of dT-22 in the active site of Tsn15, and whether the active site could accommodate the pericyclic mechanism proposed. Analogous MD simulation calculations have been performed on the [4+2] cyclase AbyU that supported its [4+2] cycloaddition mechanism[158].

Another outstanding question is how the stereochemistry of the cyclohexane and tetrahydropyran rings is established. The cyclohexane and tetrahydropyran rings of tetronasin and tetronomycin have the opposite stereochemistry. Once the mechanisms of Tsn11 and Tsn15 are known it would therefore be very interesting to compare their structures to the structures of Tmn9 and Tmn8, respectively. Such analysis has the potential to identify any key amino acid residues involved in controlling the stereochemistry at C4, C5, C10, and C13.

Following on from the last point, the substrate tolerance of Tsn11 and Tsn15 could also be explored. Analogues of T-16 and T-22 could be created synthetically or semi-synthetically and tested as substrates for Tsn11 or Tsn15, respectively. Such substrate tolerance would be a valuable feature for using Tsn11 and Tsn15 as synthetic biology tools to produce novel polyether tetronates. Now that the key of substrate-binding side chains in Tsn15 have been identified (**Chapter 6**), amino acids adjacent to the binding site could be mutated in attempt to better accommodate T-22 analogues.

Related to substrate tolerance, but unrelated to polyether tetronates, it would be interesting to determine if the tetrodecamycin biosynthetic pathway could be engineered to produce a novel “abyssomicin” compound. The biosynthesis pathway of tetrodecamycin contains TedJ, a homologue of PyrE3, that catalyses the intramolecular [4+2] cycloaddition to form the dialkyldecalin core of tetrodecamycin[165]. However, the theoretical linear tetrodecamycin precursor strongly resembles a linear abyssomicin compound. It would therefore be interesting to knock-out *tedJ* and knock-in an *abyU*-like [4+2] cyclase to determine if a different cyclisation product would be generated (**Figure 7.1**).

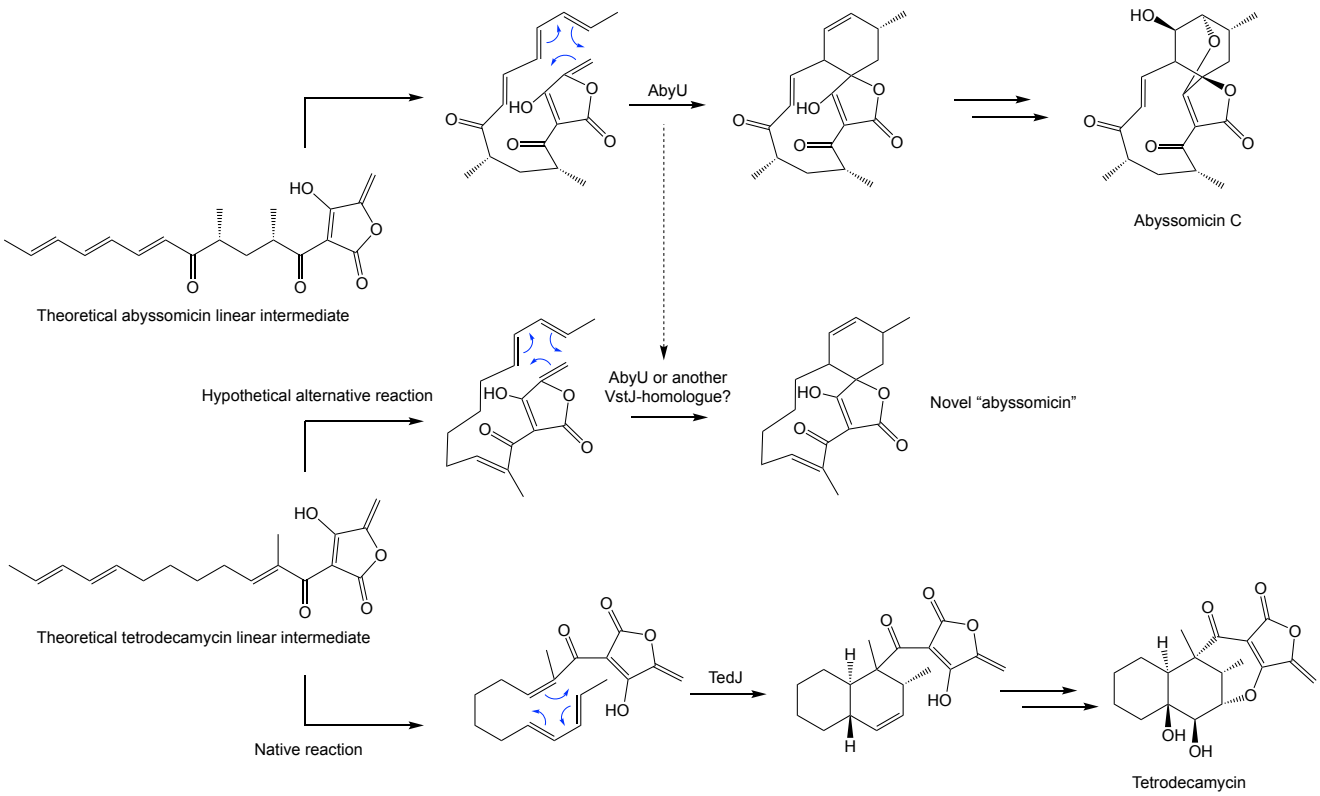


Figure 7.1 Proposed method of engineering the tetrodecamycin pathway to produce a novel abyssomicin. The *pyrE3* homologue, *tedJ*, could be deleted from the tetrodecamycin BGC and replaced with an *abyU*-like [4+2] cyclase.

A mystery that still remains in tetronasin biosynthesis is why the *tsn* cluster contains two FabH-like enzymes (Tsn9 and Tsn13). These two FabH enzymes share 45% sequence identity and previous work in the Leadlay lab demonstrated that each is essential for tetronasin biosynthesis (Sun *et al.*, *unpublished*). The relative contributions of Tsn9 and Tsn13 in catalysing tetronate ring formation (assuming that is their role) are therefore unknown. One possibility is that the two FabH-like enzymes are required for incorporating the usual glycolyl-ACP unit (rather than glyceryl-ACP), perhaps one catalysing C-C bond formation and the other catalysing C-O bond formation. FabH enzymes are homodimeric[306,307], so another possibility is that Tsn9 and Tsn13 form a heterocomplex to form the tetronate ring in a single step. The best way to test this would be if tetronate ring biosynthesis could be reconstituted *in vitro* using glycolyl-ACP, Tsn9, Tsn13, and a suitable ACP-bound β-keto thioester substrate (**Figure 7.2**) (as was done for the tetronate RK-682)[122]. Sourcing a suitable substrate would be by far the most difficult aspect of this experiment, as Tsn9 and Tsn13 would likely be substrate specific. Chemically synthesising the theoretical linear β-keto thioester intermediate of tetronasin would likely be very difficult and impede progress. Simpler β-keto thioester substrates could be tested, such as the RK-682 precursor[122], but their likelihood of being accepted by Tsn9/Tsn13 is unknown and likely to be low. An alternative strategy could be to feed the *S. longisporoflavus* Δtsn9 and *S. longisporoflavus* Δtsn13 mutants the same chemical probes used in this thesis to prematurely offload polyketide intermediates. The identity of these intermediates could provide clues as to the functions of Tsn9 and Tsn13.

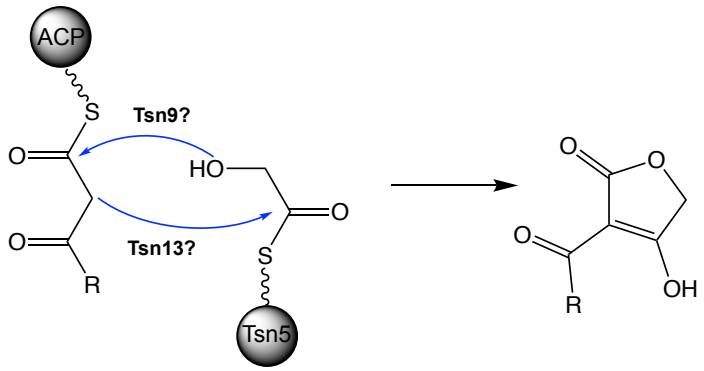


Figure 7.2 Biosynthesis of the tetronasin tetronate ring. The role of the FabH-like enzymes Tsn9 and Tsn13 in tetronasin biosynthesis is unknown, but they may act together to produce the tetronate ring.

7.4 Concluding remarks

In this work the [4+2] cyclase homologues Tsn11 and Tsn15 were shown to be responsible for forming the cyclohexane and tetrahydropyran rings of tetronasin, solving a long-standing mystery in the field of polyether tetronates. Tsn11 appears to catalyse a latent [4+2] cycloaddition reaction, forming an oxadecalin intermediate that is subsequently fragmented by Tsn15 to form tetronasin. The prevalence of such latent cyclisation reactions in other natural product biosynthesis pathways will be an interesting topic to explore further. The activity of Tsn15 also suggests that there are additional novel pericyclic reactions to be discovered in natural product biosynthesis pathways.

References

1. Borchardt, J. K. The Beginnings of Drug Therapy: Ancient Mesopotamian Medicine. *Drug News & Perspectives* **15**, 187–192 (2002).
2. Klayman, D. L. Qinghaosu (Artemisinin): An Antimalarial Drug from China. *Science* **228**, 1049–1055 (1985).
3. Krishna, S., Bustamante, L., Haynes, R. K. & Staines, H. M. Artemisinins: their Growing Importance in Medicine. *Trends in Pharmacological Sciences* **29**, 520–527 (2008).
4. Malik, V. S. Microbial Secondary Metabolism. *Trends in Biochemical Sciences* **5**, 68–72 (1980).
5. Newman, D. J. & Cragg, G. M. Natural Products as Sources of New Drugs from 1981 to 2014. *Journal of Natural Products* **79**, 629–661 (2016).
6. Harvey, A. L., Edrada-Ebel, R. & Quinn, R. J. The Re-Emergence of Natural Products for Drug Discovery in the Genomics Era. *Nature Reviews Drug Discovery* **14**, 111–129 (2015).
7. Koehn, F. E. & Carter, G. T. The Evolving Role of Natural Products in Drug Discovery. *Nature Reviews Drug Discovery* **4**, 206–220 (2005).
8. Cragg, G. M. & Newman, D. J. Natural Products: a Continuing Source of Novel Drug Leads. *Biochimica Et Biophysica Acta* **1830**, 3670–3695 (2013).
9. Fleming, A. On the Antibacterial Action of Cultures of a *Penicillium*, with Special Reference to their use in the Isolation of *B. Influenzæ*. *British Journal of Experimental Pathology* **10**, 226–236 (1929).
10. Quinn, R. Rethinking Antibiotic Research and Development: World War II and the Penicillin Collaborative. *American Journal of Public Health* **103**, 426–434 (2013).
11. Rammelkamp, C. H. & Maxon, T. Resistance of *Staphylococcus Aureus* to the Action of Penicillin. *Proceedings of the Society for Experimental Biology and Medicine* **51**, 386–389 (1942).
12. Lewis, K. Antibiotics: Recover the Lost Art of Drug Discovery. *Nature* **485**, 439–440 (2012).
13. Lyddiard, D., Jones, G. L. & Greatrex, B. W. Keeping it Simple: Lessons from the Golden Era of Antibiotic Discovery. *FEMS Microbiology Letters* **363**, (2016).

14. Zaman, S. B. *et al.* A Review on Antibiotic Resistance: Alarm Bells are Ringing. *Cureus* **9**, E1403 (2017).
15. Davies, J. & Davies, D. Origins and Evolution of Antibiotic Resistance. *Microbiology and Molecular Biology Reviews* **74**, 417–433 (2010).
16. Velayati, A. A. *et al.* Emergence of New Forms of Totally Drug-Resistant Tuberculosis Bacilli: Super Extensively Drug-Resistant Tuberculosis or Totally Drug-Resistant Strains in Iran. *Chest* **136**, 420–425 (2009).
17. Barka, E. A. *et al.* Taxonomy, Physiology, and Natural Products of Actinobacteria. *Microbiology and Molecular Biology Reviews* **80**, 1–43 (2016).
18. Chater, K. F. *Streptomyces* Inside-Out: a New Perspective on the Bacteria that Provide us with Antibiotics. *Philosophical Transactions of The Royal Society B: Biological Sciences* **361**, 761–768 (2006).
19. Bentley, S. D. *et al.* Complete Genome Sequence of the Model Actinomycete *Streptomyces coelicolor* A3(2). *Nature* **417**, 141–147 (2002).
20. Chater, K. F., Biró, S., Lee, K. J., Palmer, T. & Schrempf, H. The Complex Extracellular Biology of *Streptomyces*. *FEMS Microbiology Reviews* **34**, 171–198 (2010).
21. Haeder, S., Wirth, R., Herz, H. & Spiteller, D. Candicidin-Producing *Streptomyces* Support Leaf-Cutting Ants to Protect their Fungus Garden Against the Pathogenic Fungus *Escovopsis*. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 4742–4746 (2009).
22. Abu-Salah, K. M. Amphotericin B: An Update. *British Journal of Biomedical Science* **53**, 122–133 (1996).
23. Human, Z. R. *et al.* Antifungal *Streptomyces* Spp. Associated with the Infructescences of Protea Spp. in South Africa. *Frontiers in Microbiology* **7**, (2016).
24. Dalisay, D. S. *et al.* Marine Sediment-Derived *Streptomyces* Bacteria from British Columbia, Canada are a Promising Microbiota Resource for the Discovery of Antimicrobial Natural Products. *PLOS ONE* **8**, E77078 (2013).
25. Rajan, B. M. & Kannabiran, K. Extraction and Identification of Antibacterial Secondary Metabolites from Marine *Streptomyces* Sp. VITBRK2. *International Journal of Molecular and Cellular Medicine* **3**, 130–137 (2014).
26. Yang, A. *et al.* Nitrosporeusines A and B, Unprecedented Thioester-Bearing Alkaloids from the Arctic *Streptomyces nitrosporeus*. *Organic Letters* **15**, 5366–5369 (2013).

27. Kroiss, J. *et al.* Symbiotic Streptomycetes Provide Antibiotic Combination Prophylaxis for Wasp Offspring. *Nature Chemical Biology* **6**, 261–263 (2010).
28. Kaltenpoth, M., Göttler, W., Herzner, G. & Strohm, E. Symbiotic Bacteria Protect Wasp Larvae from Fungal Infestation. *Current Biology* **15**, 475–479 (2005).
29. Oh, D.-C., Poulsen, M., Currie, C. R. & Clardy, J. Dentigerumycin: A Bacterial Mediator of an Ant-Fungus Symbiosis. *Nature Chemical Biology* **5**, 391–393 (2009).
30. Seipke, R. F. *et al.* Fungus-Growing Allomerus Ants are Associated with Antibiotic-Producing Actinobacteria. *Antonie Van Leeuwenhoek* **101**, 443–447 (2012).
31. Scott, J. J. *et al.* Bacterial Protection of Beetle-Fungus Mutualism. *Science* **322**, 63–63 (2008).
32. Riedlinger, J. *et al.* Auxofuran, A Novel Metabolite that Stimulates the Growth of Fly Agaric, is Produced by the Mycorrhiza Helper Bacterium *Streptomyces* Strain Ach 505. *Applied and Environmental Microbiology* **72**, 3550–3557 (2006).
33. Tokala, R. K. *et al.* Novel Plant-Microbe Rhizosphere Interaction Involving *Streptomyces lydicus* WYEC108 and the Pea Plant (Pisum Sativum). *Applied and Environmental Microbiology* **68**, 2161–2171 (2002).
34. Castillo, U. F. *et al.* Munumbicins, Wide-Spectrum Antibiotics Produced by *Streptomyces* NRRL 30562, Endophytic on Kennedia Nigriscansa. *Microbiology* **148**, 2675–2685 (2002).
35. Loria, R., Kers, J. & Joshi, M. Evolution of Plant Pathogenicity in *Streptomyces*. *Annual Review of Phytopathology* **44**, 469–487 (2006).
36. Quintana, E. T. *et al. Streptomyces sudanensis* Sp. Nov., A New Pathogen Isolated from Patients with *Actinomycetoma*. *Antonie Van Leeuwenhoek* **93**, 305–313 (2008).
37. Kapadia, M., Rolston, K. V. I. & Han, X. Y. Invasive *Streptomyces* Infections: Six Cases and Literature Review. *American Journal of Clinical Pathology* **127**, 619–624 (2007).
38. Mayfield, C. I., Williams, S. T., Ruddick, S. M. & Hatfield, H. L. Studies on the Ecology of Actinomycetes in Soil IV. Observations on the Form and Growth of *Streptomycetes* in Soil. *Soil Biology and Biochemistry* **4**, 79–91 (1972).
39. Elliot, M. A. & Flärdh, K. Streptomycete Spores. in *Els* (American Cancer Society, 2012).
40. Hashimoto, T. *et al.* Biosynthesis of Quinolidomicin, the Largest Known Macrolide of Terrestrial Origin: Identification and Heterologous Expression of a Biosynthetic Gene Cluster over 200 Kb. *Organic Letters* (2018).

41. Lin, Y.-S., Kieser, H. M., Hopwood, D. A. & Chen, C. W. The Chromosomal DNA of *Streptomyces lividans* 66 is Linear. *Molecular Microbiology* **10**, 923–933 (1993).
42. Ikeda, H. *et al.* Complete Genome Sequence and Comparative Analysis of the Industrial Microorganism *Streptomyces avermitilis*. *Nature Biotechnology* **21**, 526–531 (2003).
43. Wang, X.-J. *et al.* Genome Sequence of the Milbemycin-Producing Bacterium *Streptomyces bingchenggensis*. *Journal of Bacteriology* **192**, 4526–4527 (2010).
44. Cimermancic, P. *et al.* Insights into Secondary Metabolism from a Global Analysis of Prokaryotic Biosynthetic Gene Clusters. *Cell* **158**, 412–421 (2014).
45. Lawrence, J. G. & Roth, J. R. Selfish Operons: Horizontal Transfer May Drive the Evolution of Gene Clusters. *Genetics* **143**, 1843–1860 (1996).
46. Ballouz, S., Francis, A. R., Lan, R. & Tanaka, M. M. Conditions for the Evolution of Gene Clusters in Bacterial Genomes. *PLOS Computational Biology* **6**, E1000672 (2010).
47. Omelchenko, M. V., Makarova, K. S., Wolf, Y. I., Rogozin, I. B. & Koonin, E. V. Evolution of Mosaic Operons by Horizontal Gene Transfer and Gene Displacement *in situ*. *Genome Biology* **4**, R55 (2003).
48. Pál, C. & Hurst, L. D. Evidence Against the Selfish Operon Theory. *Trends in Genetics* **20**, 232–234 (2004).
49. Fang, G., Rocha, E. P. & Danchin, A. Persistence Drives Gene Clustering in Bacterial Genomes. *BMC Genomics* **9**, 4 (2008).
50. Flores-Sanchez, I. J. & Verpoorte, R. Plant Polyketide Synthases: A Fascinating Group of Enzymes. *Plant Physiology and Biochemistry* **47**, 167–174 (2009).
51. Cox, R. J. & Simpson, T. J. Chapter 3 Fungal Type I Polyketide Synthases. in (*Methods in Enzymology*) **459**, 49–78 (Academic Press, 2009).
52. Shou, Q. *et al.* A Hybrid Polyketide-Nonribosomal Peptide in Nematodes that Promotes Larval Survival. *Nature Chemical Biology* (2016).
53. Weissman, K. J. & Leadlay, P. F. Combinatorial Biosynthesis of Reduced Polyketides. *Nature Reviews Microbiology* **3**, 925–936 (2005).
54. Rohr, J. A New Role for Polyketides. *Angewandte Chemie International Edition* **39**, 2847–2849 (2000).
55. Weissman, K. J. Chapter 1 Introduction to Polyketide Biosynthesis. in (Ed. Enzymology, B.-M. in) **459**, 3–16 (Academic Press, 2009).

56. Lee, M. D., Ellestad, G. A. & Borders, D. B. Calicheamicins: Discovery, Structure, Chemistry, and Interaction with DNA. *Accounts of Chemical Research* **24**, 235–243 (1991).
57. Staunton, J. & Weissman, K. J. Polyketide Biosynthesis: A Millennium Review. *Natural Product Reports* **18**, 380–416 (2001).
58. Bacci, L., Lupi, D., Savoldelli, S. & Rossaro, B. A Review of Spinosyns, a Derivative of Biological Acting Substances as a Class of Insecticides with a Broad Range of Action Against many Insect Pests. *Journal of Entomological & Acarological Research* **48**, (2016).
59. Watkins, S. M., Reich, A., Fleming, L. E. & Hammond, R. Neurotoxic Shellfish Poisoning. *Marine Drugs* **6**, 431–455 (2008).
60. George, K. M. *et al.* Mycolactone: A Polyketide Toxin from Mycobacterium Ulcerans Required for Virulence. *Science* **283**, 854–857 (1999).
61. Chan, Y. A., Podevels, A. M., Kevany, B. M. & Thomas, M. G. Biosynthesis of Polyketide Synthase Extender Units. *Natural Product Reports* **26**, 90–114 (2009).
62. Cronan, J. E. & Thomas, J. Bacterial Fatty Acid Synthesis and its Relationships with Polyketide Synthetic Pathways. *Methods in Enzymology* **459**, 395–433 (2009).
63. Hertweck, C., Luzhetskyy, A., Rebets, Y. & Bechthold, A. Type II Polyketide Synthases: Gaining a Deeper Insight into Enzymatic Teamwork. *Natural Product Reports* **24**, 162–190 (2007).
64. Katsuyama, Y. & Ohnishi, Y. Chapter Sixteen - Type III Polyketide Synthases in Microorganisms. in *Methods in Enzymology* (Ed. Hopwood, D. A.) **515**, 359–377 (Academic Press, 2012).
65. Keatinge-Clay, A. T. The Structures of Type I Polyketide Synthases. *Natural Product Reports* **29**, 1050–1073 (2012).
66. Weissman, K. J. Uncovering the Structures of Modular Polyketide Synthases. *Natural Product Reports* **32**, 436–453 (2015).
67. Chen, H. & Du, L. Iterative Polyketide Biosynthesis by Modular Polyketide Synthases in Bacteria. *Applied Microbiology and Biotechnology* **100**, 541–557 (2015).
68. Pöplau, P., Frank, S., Morinaka, B. I. & Piel, J. An Enzymatic Domain for the Formation of Cyclic Ethers in Complex Polyketides. *Angewandte Chemie International Edition* **52**, 13215–13218 (2013).

69. Berkhan, G. & Hahn, F. A Dehydratase Domain in Ambruticin Biosynthesis Displays Additional Activity as a Pyran-Forming Cyclase. *Angewandte Chemie International Edition* **53**, 14240–14244 (2014).
70. Ma, M., Lohman, J. R., Liu, T. & Shen, B. C-S Bond Cleavage by a Polyketide Synthase Domain. *Proceedings of the National Academy of Sciences of the United States of America* **112**, 10359–10364 (2015).
71. Sundaram, S., Heine, D. & Hertweck, C. Polyketide Synthase Chimeras Reveal Key Role of Ketosynthase Domain in Chain Branching. *Nature Chemical Biology* **11**, 949-51 (2015).
72. Helfrich, E. J. N. & Piel, J. Biosynthesis of Polyketides by Trans-AT Polyketide Synthases. *Natural Product Reports* **33**, 231–316 (2016).
73. Piel, J. A Polyketide Synthase-Peptide Synthetase Gene Cluster from an Uncultured Bacterial Symbiont of Paederus Beetles. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 14002–14007 (2002).
74. Donadio, S., Staver, M. J., Mcalpine, J. B., Swanson, S. J. & Katz, L. Modular Organization of Genes Required for Complex Polyketide Biosynthesis. *Science* **252**, 675–679 (1991).
75. Cortés, J., Haydock, S. F., Roberts, G. A., Bevitt, D. J. & Leadlay, P. F. An Unusually Large Multifunctional Polypeptide in the Erythromycin-Producing Polyketide Synthase of *Saccharopolyspora erythraea*. *Nature* **348**, 176–178 (1990).
76. Bevitt, D. J., Cortes, J., Haydock, S. F. & Leadlay, P. F. 6-Deoxyerythronolide-B Synthase 2 from *Saccharopolyspora erythraea*. *European Journal of Biochemistry* **204**, 39–49 (1992).
77. Tuan, J. S. *et al.* Cloning of Genes Involved in Erythromycin Biosynthesis from *Saccharopolyspora erythraea* using a Novel Actinomycete-*Escherichia coli* Cosmid. *Gene* **90**, 21–29 (1990).
78. Taubman, S. B., Young, F. E. & Corcoran, J. W. Antibiotic Glycosides, Iv. Studies on the Mechanism of Erythromycin Resistance in Bacillus Subtilis. *Proceedings of the National Academy of Sciences of the United States of America* **50**, 955–962 (1963).
79. Taubman, S. B., Jones, N. R., Young, F. E. & Corcoran, J. W. Sensitivity and Resistance to Erythromycin in Bacillus Subtilis 168: The Ribosomal Binding of Erythromycin and Chloramphenicol. *Biochimica Et Biophysica Acta (BBA) - Nucleic Acids and Protein Synthesis* **123**, 438–440 (1966).

80. Long, P. F. *et al.* Engineering Specificity of Starter Unit Selection by the Erythromycin-Producing Polyketide Synthase. *Molecular Microbiology* **43**, 1215–1225 (2002).
81. Bisang, C. *et al.* A Chain Initiation Factor Common to Both Modular and Aromatic Polyketide Synthases. *Nature* **401**, 502–505 (1999).
82. Witkowski, A., Joshi, A. K., Lindqvist, Y. & Smith, S. Conversion of a B-Ketoacyl Synthase to a Malonyl Decarboxylase by Replacement of the Active-Site Cysteine with Glutamine. *Biochemistry* **38**, 11643–11650 (1999).
83. Weissman, K. J. *et al.* The Molecular Basis of Celmer's Rules: The Stereochemistry of the Condensation Step in Chain Extension on the Erythromycin Polyketide Synthase. *Biochemistry* **36**, 13849–13855 (1997).
84. Cane, D. E., Liang, T. C., Taylor, P. B., Chang, C. & Yang, C. C. Macrolide Biosynthesis. 3. Stereochemistry of the Chain-Elongation Steps of Erythromycin Biosynthesis. *Journal of the American Chemical Society* **108**, 4957–4964 (1986).
85. Broadhurst, R. W., Nietlispach, D., Wheatcroft, M. P., Leadlay, P. F. & Weissman, K. J. The Structure of Docking Domains in Modular Polyketide Synthases. *Chemistry & Biology* **10**, 723–731 (2003).
86. Staunton, J. & Wilkinson, B. Biosynthesis of Erythromycin and Rapamycin. *Chemical Reviews* **97**, 2611–2630 (1997).
87. Marsden, A. F. *et al.* Stereospecific Acyl Transfers on the Erythromycin-Producing Polyketide Synthase. *Science* **263**, 378–380 (1994).
88. Khayatt, B. I., Overmars, L., Siezen, R. J. & Francke, C. Classification of the Adenylation and Acyl-Transferase Activity of NRPS and PKS Systems using Ensembles of Substrate Specific Hidden Markov Models. *PLOS ONE* **8**, E62136 (2013).
89. Bonnett, S. A. *et al.* Acyl-Coa Subunit Selectivity in the Pikromycin Polyketide Synthase Pikaiv: Steady-State Kinetics and Active-Site Occupancy Analysis by FTICR-MS. *Chemistry & Biology* **18**, 1075–1081 (2011).
90. Tang, Y., Kim, C.-Y., Mathews, I. I., Cane, D. E. & Khosla, C. The 2.7-Å Crystal Structure of a 194-Kda Homodimeric Fragment of the 6-Deoxyerythronolide B Synthase. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 11124–11129 (2006).
91. Haydock, S. F. *et al.* Divergent Sequence Motifs Correlated with the Substrate Specificity of (Methyl)Malonyl-Coa:Acyl Carrier Protein Transacylase Domains in Modular Polyketide Synthases. *FEBS Letters* **374**, 246–248 (1995).

92. Del Vecchio, F. *et al.* Active-Site Residue, Domain and Module Swaps in Modular Polyketide Synthases. *Journal of Industrial Microbiology & Biotechnology* **30**, 489–494 (2003).
93. Beld, J., Sonnenschein, E. C., Vickery, C. R., Noel, J. P. & Burkart, M. D. The Phosphopantetheinyl Transferases: Catalysis of a Posttranslational Modification Crucial for Life. *Natural Product Reports* **31**, 61–108 (2014).
94. Yasgar, A. *et al.* A Strategy to Discover Inhibitors of Bacillus Subtilis Surfactin-Type Phosphopantetheinyl Transferase. *Molecular Biosystems* **6**, 365–375 (2010).
95. Kosa, N. M., Foley, T. L. & Burkart, M. D. Fluorescent Techniques for Discovery and Characterization of Phosphopantetheinyl Transferase Inhibitors. *The Journal of Antibiotics* **67**, 113–120 (2014).
96. Owen, J. G., Copp, J. N. & Ackerley, D. F. Rapid and Flexible Biochemical Assays for Evaluating 4'-Phosphopantetheinyl Transferase Activity. *Biochemical Journal* **436**, 709–717 (2011).
97. Moretto, L., Vance, S., Heames, B. & Broadhurst, R. W. Dissecting how Modular Polyketide Synthase Ketoreductases Interact with Acyl Carrier Protein-Attached Substrates. *Chemical Communications* **53**, 11457–11460 (2017).
98. Aparicio, J. F. *et al.* Organization of the Biosynthetic Gene Cluster for Rapamycin in *Streptomyces hygroscopicus*: Analysis of the Enzymatic Domains in the Modular Polyketide Synthase. *Gene* **169**, 9–16 (1996).
99. Robbins, T., Kapilivsky, J., Cane, D. E. & Khosla, C. Roles of Conserved Active Site Residues in the Ketosynthase Domain of an Assembly Line Polyketide Synthase. *Biochemistry* **55**, 4476–4484 (2016).
100. Watanabe, K., Wang, C. C. C., Boddy, C. N., Cane, D. E. & Khosla, C. Understanding Substrate Specificity of Polyketide Synthase Modules by Generating Hybrid Multimodular Synthases. *Journal of Biological Chemistry* **278**, 42020–42026 (2003).
101. Keatinge-Clay, A. T. A Tylosin Ketoreductase Reveals How Chirality is Determined in Polyketides. *Chemistry & Biology* **14**, 898–908 (2007).
102. Palaniappan, N., Alhamadsheh, M. M. & Reynolds, K. A. Cis-Δ2,3-Double Bond of Phoslactomycins is Generated by a Post-PKS Tailoring Enzyme. *Journal of the American Chemical Society* **130**, 12236–12237 (2008).
103. He, H.-Y., Tang, M.-C., Zhang, F. & Tang, G.-L. Cis-Double Bond Formation by Thioesterase and Transfer by Ketosynthase in FR901464 Biosynthesis. *Journal of the American Chemical Society* **136**, 4488–4491 (2014).

104. Alhamadsheh, M. M., Palaniappan, N., Daschouduri, S. & Reynolds, K. A. Modular Polyketide Synthases and Cis-Double Bond Formation: Establishment of Activated *cis*-3-Cyclohexylpropenoic Acid as the Diketide Intermediate in Phoslactomycin Biosynthesis. *Journal of the American Chemical Society* **129**, 1910–1911 (2007).
105. Kwan, D. H. *et al.* Prediction and Manipulation of the Stereochemistry of Enoylreduction in Modular Polyketide Synthases. *Chemistry & Biology* **15**, 1231–1240 (2008).
106. Kwan, D. H. & Leadlay, P. F. Mutagenesis of a Modular Polyketide Synthase Enoylreductase Domain Reveals Insights into Catalysis and Stereospecificity. *ACS Chemical Biology* **5**, 829–838 (2010).
107. Gerber, R., Lou, L. & Du, L. A PLP-Dependent Polyketide Chain Releasing Mechanism in the Biosynthesis of Mycotoxin Fumonisins in Fusarium Verticillioides. *Journal of the American Chemical Society* **131**, 3148–3149 (2009).
108. Harvey, B. M. *et al.* Evidence that a Novel Thioesterase is Responsible for Polyketide Chain Release During Biosynthesis of the Polyether Ionophore Monensin. *Chembiochem* **7**, 1435–1442 (2006).
109. Vieweg, L., Reichau, S., Schobert, R., Leadlay, P. F. & Süssmuth, R. D. Recent Advances in the Field of Bioactive Tetronates. *Natural Product Reports* **31**, 1554–1584 (2014).
110. Gottardi, E. M. *et al.* Abyssomicin Biosynthesis: Formation of an Unusual Polyketide, Antibiotic-Feeding Studies and Genetic Analysis. *Chembiochem* **12**, 1401–1410 (2011).
111. Demydchuk, Y. *et al.* Analysis of the Tetronomycin Gene Cluster: Insights into the Biosynthesis of a Polyether Tetronate Antibiotic. *Chembiochem* **9**, 1136–1145 (2008).
112. Daduang, R. *et al.* Characterization of the Biosynthetic Gene Cluster for Maklamicin, a Spirotetronate-Class Antibiotic of the Endophytic Micromonospora Sp. NBRC 110955. *Microbiological Research* **180**, 30–39 (2015).
113. Fang, J. *et al.* Cloning and Characterization of the Tetrocarcin A Gene Cluster from Micromonospora Chalcea NRRL 11289 Reveals a Highly Conserved Strategy for Tetronate Biosynthesis in Spirotetronate Antibiotics. *Journal of Bacteriology* **190**, 6014–6025 (2008).
114. Li, S. *et al.* Dissecting Glycosylation Steps in Lobophorin Biosynthesis Implies an Iterative Glycosyltransferase. *Organic Letters* **15**, 1374–1377 (2013).

115. Zhang, H. *et al.* Elucidation of the Kijanimicin Gene Cluster: Insights into the Biosynthesis of Spirotetronate Antibiotics and Nitrosugars. *Journal of the American Chemical Society* **129**, 14670–14683 (2007).
116. Jia, X.-Y. *et al.* Genetic Characterization of the Chlorothricin Gene Cluster as a Model for Spirotetronate Antibiotic Biosynthesis. *Chemistry & Biology* **13**, 575–585 (2006).
117. He, H.-Y. *et al.* Quartromicin Biosynthesis: two Alternative Polyketide Chains Produced by One Polyketide Synthase Assembly Line. *Chemistry & Biology* **19**, 1313–1323 (2012).
118. Gverzdys, T., Kramer, G. & Nodwell, J. R. Tetrodecamycin: an Unusual and Interesting Tetronate Antibiotic. *Bioorganic & Medicinal Chemistry* **24**, 6269–6275 (2016).
119. Clutterbuck, P. W., Raistrick, H. & Reuter, F. Studies in the Biochemistry of Micro-Organisms. *Biochemical Journal* **29**, 1300–1309 (1935).
120. Keller‐Schierlein, W., Muntwyler, R., Pache, W. & Zähner, H. Stoffwechselprodukte Von Mikroorganismen 73. Mitteilung [1] Chlorothricin Und Des-Chlorothoricin. *Helvetica Chimica Acta* **52**, 127–142 (1969).
121. Sun, Y., Hong, H., Gillies, F., Spencer, J. B. & Leadlay, P. F. Glyceryl-S-Acyl Carrier Protein as an Intermediate in the Biosynthesis of Tetronate Antibiotics. *Chembiochem* **9**, 150–156 (2008).
122. Sun, Y. *et al. in vitro* Reconstruction of Tetronate RK-682 Biosynthesis. *Nature Chemical Biology* **6**, 99–101 (2010).
123. Kanchanabanca, C. *et al.* Unusual Acetylation–Elimination in the Formation of Tetronate Antibiotics. *Angewandte Chemie International Edition* **52**, 5785–5788 (2013).
124. Diels, O. & Alder, K. Synthesen in Der Hydroaromatischen Reihe. *Justus Liebigs Annalen Der Chemie* **460**, 98–122 (1928).
125. Fleming, I. *Pericyclic Reactions*. (Oxford University Press, 2015).
126. Nicolaou, K. C., Snyder, S. A., Montagnon, T. & Vassilikogiannakis, G. The Diels–Alder Reaction in Total Synthesis. *Angewandte Chemie International Edition* **41**, 1668–1698 (2002).
127. Fernández, I. & Bickelhaupt, F. M. Origin of the "*Endo* Rule" in Diels–Alder Reactions. *Journal of Computational Chemistry* **35**, 371–376 (2014).

128. Hilvert, D., Hill, K. W., Nared, K. D. & Auditor, M. T. M. Antibody Catalysis of the Diels-Alder Reaction. *Journal of the American Chemical Society* **111**, 9261–9262 (1989).
129. Braisted, A. C. & Schultz, P. G. An Antibody-Catalyzed Bimolecular Diels-Alder Reaction. *Journal of the American Chemical Society* **112**, 7430–7431 (1990).
130. Gouverneur, V. E. *et al.* Control of the *Exo* and *Endo* Pathways of the Diels-Alder Reaction by Antibody Catalysis. *Science* **262**, 204–208 (1993).
131. Meekel, A. A. P., Resmini, M. & Pandit, U. K. First Example of an Antibody-Catalysed Hetero-Diels-Alder Reaction. *Journal of the Chemical Society, Chemical Communications* 571–572 (1995).
132. Shi, Z.-D. *et al.* First Example of an Antibody-Catalyzed Aza Diels–Alder Reaction. *Bioorganic & Medicinal Chemistry Letters* **12**, 2321–2324 (2002).
133. Seelig, B. & Jäschke, A. A Small Catalytic RNA Motif with Diels-Alderase Activity. *Chemistry & Biology* **6**, 167–176 (1999).
134. Hilvert, D. Critical Analysis of Antibody Catalysis. *Annual Review of Biochemistry* **69**, 751–793 (2000).
135. Siegel, J. B. *et al.* Computational Design of an Enzyme Catalyst for a Stereoselective Bimolecular Diels-Alder Reaction. *Science* **329**, 309–313 (2010).
136. Lichman, B. R., O'Connor, S. E. & Kries, H. Biocatalytic Strategies Towards [4+2] Cycloadditions. *Chemistry–A European Journal* (2019).
137. Jeon, B., Wang, S.-A., Ruszczycky, M. W. & Liu, H. Natural [4+2]-Cyclases. *Chemical Reviews* **117**, 5367–5388 (2017).
138. Klas, K., Tsukamoto, S., Sherman, D. H. & Williams, R. M. Natural Diels–Alderases: Elusive and Irresistable. *The Journal of Organic Chemistry* **80**, 11672–11685 (2015).
139. Oikawa, H., Suzuki, Y., Naya, A., Katayama, K. & Ichihara, A. First Direct Evidence in Biological Diels-Alder Reaction of Incorporation of Diene-Dienophile Precursors in the Biosynthesis of Solanapyrones. *Journal of The American Chemical Society* **116**, 3605–3606 (1994).
140. Mizushina, Y. *et al.* A Plant Phytotoxin, Solanapyrone A, is an Inhibitor of DNA Polymerase B and Λ. *Journal of Biological Chemistry* **277**, 630–638 (2002).
141. Oikawa, H., Katayama, K., Suzuki, Y. & Ichihara, A. Enzymatic Activity Catalysing Exo-Selective Diels–Alder Reaction in Solanapyrone Biosynthesis. *Journal of the Chemical Society, Chemical Communications* **0**, 1321–1322 (1995).

142. Kasahara, K. *et al.* Solanapyrone Synthase, A Possible Diels–Alderase and Iterative Type I Polyketide Synthase Encoded in a Biosynthetic Gene Cluster from Alternaria Solani. *Chembiochem* **11**, 1245–1252 (2010).
143. Oikawa, H., Kobayashi, T., Katayama, K., Suzuki, Y. & Ichihara, A. Total Synthesis of (−)-Solanapyrone A via Enzymatic Diels−Alder Reaction of Prosolanapyrone. *The Journal of Organic Chemistry* **63**, 8748–8756 (1998).
144. Tobert, J. A. Lovastatin and Beyond: the History of the HMG-Coa Reductase Inhibitors. *Nature Reviews Drug Discovery* **2**, 517 (2003).
145. Hendrickson, L. *et al.* Lovastatin Biosynthesis in Aspergillus Terreus: Characterization of Blocked Mutants, Enzyme Activities and a Multifunctional Polyketide Synthase Gene. *Chemistry & Biology* **6**, 429–439 (1999).
146. Auclair, K. *et al.* Lovastatin Nonaketide Synthase Catalyzes an Intramolecular Diels−Alder Reaction of a Substrate Analogue. *Journal of the American Chemical Society* **122**, 11519–11520 (2000).
147. Xu, W. *et al.* Lovg: The Thioesterase Required for Dihydromonacolin L Release and Lovastatin Nonaketide Synthase Turnover in Lovastatin Biosynthesis. *Angewandte Chemie International Edition* **52**, 6472–6475 (2013).
148. Witter, D. J. & Vederas, J. C. Putative Diels−Alder-Catalyzed Cyclization During the Biosynthesis of Lovastatin. *The Journal of Organic Chemistry* **61**, 2613–2623 (1996).
149. Kim, H. J., Ruszczycky, M. W., Choi, S., Liu, Y. & Liu, H. Enzyme-Catalysed [4+2] Cycloaddition is a Key Step in the Biosynthesis of Spinosyn A. *Nature* **473**, 109–112 (2011).
150. Kirst, H. A. The Spinosyn Family of Insecticides: Realizing the Potential of Natural Products Research. *The Journal of Antibiotics* **63**, 101–111 (2010).
151. Waldron, C. *et al.* Cloning and Analysis of the Spinosad Biosynthetic Gene Cluster of *Saccharopolyspora Spinosa*. *Chemistry & Biology* **8**, 487–499 (2001).
152. Mertz, F. P. & Yao, R. C. *Saccharopolyspora Spinosa* sp. Nov. Isolated from Soil Collected in a Sugar Mill Rum Still. *International Journal of Systematic and Evolutionary Microbiology* **40**, 34–39 (1990).
153. Kim, H. J., Pongdee, R., Wu, Q., Hong, L. & Liu, H. The Biosynthesis of Spinosyn in *Saccharopolyspora Spinosa*: Synthesis of the Cross-Bridging Precursor and Identification of the Function of SpnJ. *Journal of the American Chemical Society* **129**, 14582–14584 (2007).

154. Chen, N., Zhang, F., Wu, R. & Hess, B. A. Biosynthesis of Spinosyn A: A [4+2] or [6+4] Cycloaddition? *ACS Catalysis* **8**, 2353–2358 (2018).
155. Fage, C. D. *et al.* The Structure of Spnf, a Standalone Enzyme that Catalyzes [4+2] Cycloaddition. *Nature Chemical Biology* **11**, 256–258 (2015).
156. Park, H.-R., Furihata, K., Hayakawa, Y. & Shin-Ya, K. Versipelostatin, a Novel GRP78/Bip Molecular Chaperone Down-Regulator of Microbial Origin. *Tetrahedron Letters* **43**, 6941–6945 (2002).
157. Hashimoto, T. *et al.* Biosynthesis of Versipelostatin: Identification of an Enzyme-Catalyzed [4+2]-Cycloaddition Required for Macrocyclization of Spirotetronate-Containing Polyketides. *Journal of the American Chemical Society* **137**, 572–575 (2015).
158. Byrne, M. J. *et al.* The Catalytic Mechanism of a Natural Diels–Alderase Revealed in Molecular Detail. *Journal of the American Chemical Society* (2016).
159. Zheng, Q. *et al.* Enzyme-Dependent [4+2] Cycloaddition Depends on Lid-Like Interaction of the *N*-Terminal Sequence with the Catalytic Core in Pyri4. *Cell Chemical Biology* **23**, 352–360 (2016).
160. Hofmann, E., Zerbe, P. & Schaller, F. The Crystal Structure of *Arabidopsis Thaliana* Allene Oxide Cyclase: Insights into the Oxylipin Cyclization Reaction. *Plant Cell* **18**, 3201–3217 (2006).
161. Tian, Z. *et al.* An Enzymatic [4+2] Cyclization Cascade Creates the Pentacyclic Core of Pyrroindomycins. *Nature Chemical Biology* **11**, 259–265 (2015).
162. Ding, W. *et al.* Pyrroindomycins, Novel Antibiotics Produced by *Streptomyces rugosporus* Sp. LL-42D005. *The Journal of Antibiotics* **47**, 1250–1257 (1994).
163. Zheng, Q. *et al.* Structural Insights into a Flavin-Dependent [4+2] Cyclase that Catalyzes *trans*-Decalin Formation in Pyrroindomycin Biosynthesis. *Cell Chemical Biology* **25**, 718-727.E3 (2018).
164. Tsuchida, T. *et al.* Tetrodecamycin, a New Antimicrobial Antibiotic from *Streptomyces*. *The Journal of Antibiotics* **47**, 386–388 (1994).
165. Gverzdys, T. & Nodwell, J. R. The Biosynthetic Genes for the Tetrodecamycin Antibiotics. *Journal of Bacteriology* **198**, 1965–1973 (2016).
166. Keller-Juslén, C. *et al.* Tetronomycin, a Novel Polyether of Unusual Structure. *The Journal of Antibiotics* **35**, 142–150 (1982).
167. Davies, D. H., Snape, E. W., Suter, P. J., King, T. J. & Falshaw, C. P. Structure of Antibiotic M139603; X-Ray Crystal Structure of the 4-Bromo-3, 5-Dinitrobenzoyl

Derivative. *Journal of the Chemical Society, Chemical Communications* 1073–1074 (1981).

168. Hatsu, M. *et al.* SF2487, a New Polyether Antibiotic Produced by *Actinomadura*. *The Journal of Antibiotics* **43**, 259–266 (1990).
169. Hamill, R. L. & Yao, R. C.-F. Antibiotic A80577 and Process for its Production. (1990).
170. Igarashi, Y. *et al.* Nonthmicin, a Polyether Polyketide Bearing a Halogen-Modified Tetronate with Neuroprotective and Antiinvasive Activity from Actinomadura Sp. *Organic Letters* **19**, 1406–1409 (2017).
171. Wyche, T. P. *et al.* Chemical Genomics, Structure Elucidation, and *in vivo* Studies of the Marine-Derived Anticlostridial Ecteinamycin. *ACS Chemical Biology* (2017).
172. Rutkowski, J. & Brzezinski, B. Structures and Properties of Naturally Occurring Polyether Antibiotics. *Biomed Research International* (2013).
173. Naujokat, C. & Steinhart, R. Salinomycin as a Drug for Targeting Human Cancer Stem Cells. *Journal of Biomedicine and Biotechnology* **2012**, (2012).
174. Bulsing, J. M. *et al.* Biosynthesis of the Polyketide Antibiotic ICI139603 in *Streptomyces longisporoflavus*: Assignment of the ^{13}C N.M.R. Spectrum by Two-Dimensional Methods, and Determination of the Origin of the Carbon Atoms. *Journal of the Chemical Society, Chemical Communications* **19**, 1301–1302 (1984).
175. Demetriadou, A. K. *et al.* Biosynthesis of the Polyketide Polyether Antibiotic ICI 139603 in *Streptomyces longisporoflavus* from ^{18}O-Labelled Acetate and Propionate. *Journal of the Chemical Society, Chemical Communications* **7**, 408–410 (1985).
176. Doddrell, D. M. *et al.* Biosynthesis of the Polyether Antibiotic ICI139603 in *Streptomyces longisporoflavus*: Investigation of Deuterium Retention After Incorporation of $CD_3{}^{13}CO_2H$, $^{13}CD_3CO_2H$, and $CH_3CD_2{}^{13}CO_2H$ using ^{2}H N.M.R. and Edited ^{13}C N.M.R. Spectra. *Journal of the Chemical Society, Chemical Communications* **19**, 1302–1304 (1984).
177. Hailes, H. C., Jackson, C. M., Leadlay, P. F., Ley, S. V. & Staunton, J. Biosynthesis of Tetronasin: Part 1 Introduction and Investigation of the Diketide and Triketide Intermediates Bound to the Polyketide Synthase. *Tetrahedron Letters* **35**, 307–310 (1994).
178. Agtarap, Amelia., Chamberlin, J. W., Pinkerton, M. & Steinrauf, L. K. Structure of Monensic Acid, a New Biologically Active Compound. *Journal of the American Chemical Society* **89**, 5737–5739 (1967).

179. Miyazaki, Y. *et al.* Salinomycin, a New Polyether Antibiotic. *The Journal of Antibiotics* **27**, 814–821 (1974).
180. Ley, S. V. *et al.* Synthesis of the Acyltetronic Acid Ionophore Tetronasin (ICI M139603). *Journal of the Chemical Society 1* 2259–2276 (1998).
181. Hori, K., Kazuno, H., Nomura, K. & Yoshii, E. The First Total Synthesis of Tetronasin (M139603). *Tetrahedron Letters* **34**, 2183–2186 (1993).
182. Boons, G.-J. *et al.* Novel Polyene Cyclisation Routes to the Acyl Tetronic Acid Ionophore Tetronasin (ICI M139603). *Tetrahedron Letters* **35**, 323–326 (1994).
183. Ley, S. V., Maw, G. N. & Trudell, M. L. Enantioselective Synthesis of the C3-C11 Hydrocarbon Fragment of the Ionophore Antibiotic Tetronasin (Ici 139603). *Tetrahedron Letters* **31**, 5521–5524 (1990).
184. Hyo, W. L., Ihl-Young, C. L. & Sung, K. K. Synthesis of the Tetrahydrofuran Fragment of Tetronasin. *Tetrahedron Letters* **31**, 7637–7640 (1990).
185. Grandjean, J. & Laszlo, P. Synergistic Transport of Praseodymium^{3+} Ion Across Lipid Bilayers in the Presence of two Chemically Distinct Ionophores. *Journal of the American Chemical Society* **106**, 1472–1476 (1984).
186. Riddell, F. G. & Arumugam, S. The Transport of $Li^{+,}$ Na^{+} and K^{+} Ions Through Phospholipid Bilayers Mediated by the Antibiotic M139603 Studied By7li-, 23Na- and 39K-NMR. *Biochimica Et Biophysica Acta (BBA) - Biomembranes* **984**, 6–10 (1989).
187. Newbold, C. J., Wallace, R. J., Watt, N. D. & Richardson, A. J. Effect of the Novel Ionophore Tetronasin (ICI 139603) on Ruminal Microorganisms. *Applied and Environmental Microbiology* **54**, 544–547 (1988).
188. Gates, R. N., Roland, L. T., Wyatt, W. E., Hembry, F. G. & Bailie, J. H. Dose-Response Relationship of Tetronasin Administered to Grazing Steers. *Journal of Animal Science* **67**, 3419–3424 (1989).
189. Aitchison, E. M., Tanaka, K. & Rowe, J. B. Evaluation of Feed Additives for Increasing Wool Production from Merino Sheep. 2. Flavomycin and Tetronasin Included in Lucerne-Based Pellets or Wheaten Chaff Fed Ad Libitum. *Australian Journal of Experimental Agriculture* **29**, 327–332 (1989).
190. Cane, D. E., Liang, T. C. & Hasler, H. Polyether Biosynthesis. 2. Origin of the Oxygen Atoms of Monensin A. *Journal of the American Chemical Society* **104**, 7274–7281 (1982).

191. Hemmerling, F. & Hahn, F. Biosynthesis of Oxygen and Nitrogen-Containing Heterocycles in Polyketides. *Beilstein Journal of Organic Chemistry* **12**, 1512–1550 (2016).
192. Bhatt, A. *et al.* Accumulation of an *E,E,E*-Triene by the Monensin-Producing Polyketide Synthase When Oxidative Cyclization is Blocked. *Angewandte Chemie* **117**, 7237–7240 (2005).
193. Gallimore, A. R. *et al.* Evidence for the Role of the Monb Genes in Polyether Ring Formation During Monensin Biosynthesis. *Chemistry & Biology* **13**, 453–460 (2006).
194. Shichijo, Y. *et al.* Epoxide Hydrolase Lsd19 for Polyether Formation in the Biosynthesis of Lasalocid A: Direct Experimental Evidence on Polyene-Polyepoxide Hypothesis in Polyether Biosynthesis. *Journal of the American Chemical Society* **130**, 12230–12231 (2008).
195. Hailes, H. C. *et al.* Biosynthesis of Tetronasin: Part 2 Identification of the Tetraketide Intermediate Attached to the Polyketide Synthase. *Tetrahedron Letters* **35**, 311–314 (1994).
196. Tosin, M., Smith, L. & Leadlay, P. F. Insights into Lasalocid A Ring Formation by Chemical Chain Termination *in vivo*. *Angewandte Chemie International Edition Engl.* **50**, 11930–11933 (2011).
197. Luhavaya, H. *et al.* Enzymology of Pyran Ring A Formation in Salinomycin Biosynthesis. *Angewandte Chemie* **127**, 13826–13829 (2015).
198. Linton, K. J., Cooper, H. N., Hunter, L. S. & Leadlay, P. F. An ABC-Transporter from *Streptomyces longisporoflavus* Confers Resistance to the Polyether-Ionophore Antibiotic Tetronasin. *Molecular Microbiology* **11**, 777–785 (1994).
199. Horn, H., Cheng, C., Edrada-Ebel, R., Hentschel, U. & Abdelmohsen, U. R. Draft Genome Sequences of three Chemically Rich Actinomycetes Isolated from Mediterranean Sponges. *Marine Genomics* **24**, 285–287 (2015).
200. Benaud, N. *et al.* Harnessing Long-Read Amplicon Sequencing to Uncover NRPS and Type I PKS Gene Sequence Diversity in Polar Desert Soils. *FEMS Microbiol Ecology*
201. Grandjean, J. & Laszlo, P. Solution Structure and Cation-Binding Abilities of two Quasi-Isomorphous Antibiotic Ionophores, M 139603 and Tetronomycin. *Tetrahedron Letters* **24**, 3319–3322 (1983).
202. Semmelhack, M. F., Kim, C. R., Dobler, W. & Meier, M. Controlled β-Hydride Elimination During Tetrahydropyran Formation with Pd(II); Diastereoselective

Formation of the Tetrahydropyran Ring of Tetronomycin. *Tetrahedron Letters* **30**, 4925–4928 (1989).

203. Semmelhack, M. F. *et al.* Palladium-Promoted Synthesis of Ionophore Antibiotics. Strategy and Assembly of the Homochiral Tetrahydrofuran and Tetrahydropyran Portions of Tetronomycin. *Journal of the American Chemical Society* **116**, 7455–7456 (1994).
204. Hori, K. *et al.* Total Synthesis of Tetronomycin. *The Journal of Organic Chemistry* **57**, 2888–2902 (1992).
205. Hori, K., Nomura, K., Mori, S. & Yoshii, E. Synthesis of the Acyltetronic Acid Fragment of Tetronomycin. *Journal of the Chemical Society, Chemical Communications* **11**, 712–713 (1989).
206. Wilkinson, C. J. *et al.* Increasing the Efficiency of Heterologous Promoters in Actinomycetes. *Journal of Molecular Microbiology and Biotechnology* **4**, 417–426 (2002).
207. Macneil, D. J. *et al.* Analysis of *Streptomyces avermitilis* Genes Required for Avermectin Biosynthesis Utilizing a Novel Integration Vector. *Gene* **111**, 61–68 (1992).
208. Less, S. Biosynthetic Studies on Tetronasin. (University of Cambridge, 1996).
209. Jackson, C. M. Tetronasin Biosynthesis. (University of Cambridge, 1988).
210. Kieser, T., Bibb, M. J., Buttner, M. J., Chater, K. F. & Hopwood, D. A. *Practical Streptomyces Genetics*. (John Innes Foundation, 2000).
211. Gibson, D. G. *et al.* Enzymatic Assembly of DNA Molecules up to Several Hundred Kilobases. *Nature Methods* **6**, 343–345 (2009).
212. Fu, C., Donovan, W. P., Shikapwashya-Hasser, O., Ye, X. & Cole, R. H. Hot Fusion: an Efficient Method to Clone Multiple DNA Fragments as well as Inverted Repeats without Ligase. *PLOS ONE* **9**, E115318 (2014).
213. Harding, S. E. & Horton, J. C. *Analytical Ultracentrifugation in Biochemistry and Polymer Science*. (Royal Society of Chemistry, 1992).
214. Kabsch, W. XDS. *Acta Crystallographica. Section D, Biological Crystallography* **66**, 125–132 (2010).
215. Evans, P. R. & Murshudov, G. N. How Good are my Data and What is the Resolution? *Acta Crystallographica. Section D, Biological Crystallography* **69**, 1204–1214 (2013).

216. Winn, M. D. *et al.* Overview of the CCP4 Suite and Current Developments. *Acta Crystallographica. Section D, Biological Crystallography* **67**, 235–242 (2011).
217. Terwilliger, T. C. *et al.* Decision-Making in Structure Solution using Bayesian Estimates of Map Quality: the PHENIX Autosol Wizard. *Acta Crystallographica. Section D, Biological Crystallography* **65**, 582–601 (2009).
218. Terwilliger, T. C. *et al.* Iterative Model Building, Structure Refinement and Density Modification with the PHENIX Autobuild Wizard. *Acta Cryst D* **64**, 61–69 (2008).
219. Afonine, P. V. *et al.* Towards Automated Crystallographic Structure Refinement with Phenix.Refine. *Acta Crystallographica. Section D, Biological Crystallography* **68**, 352–367 (2012).
220. Emsley, P. & Cowtan, K. Coot: Model-Building Tools for Molecular Graphics. *Acta Crystallographica. Section D, Biological Crystallography* **60**, 2126–2132 (2004).
221. Chen, V. B. *et al.* Molprobity: All-Atom Structure Validation for Macromolecular Crystallography. *Acta Crystallographica. Section D, Biological Crystallography* **66**, 12–21 (2010).
222. Mccoy, A. J. Solving Structures of Protein Complexes by Molecular Replacement with Phaser. *Acta Crystallographica. Section D, Biological Crystallography* **63**, 32–41 (2007).
223. Adams, P. D. *et al.* The Phenix Software for Automated Determination of Macromolecular Structures. *Methods* **55**, 94–106 (2011).
224. Le, S. Q. & Gascuel, O. An Improved General Amino Acid Replacement Matrix. *Molecular Biology and Evolution* **25**, 1307–1320 (2008).
225. Kumar, S., Stecher, G. & Tamura, K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Molecular Biology and Evolution* **33**, 1870–1874 (2016).
226. Hall, B. G. Building Phylogenetic Trees from Molecular Data with MEGA. *Molecular Biology and Evolution* **30**, 1229–1235 (2013).
227. Sievers, F. *et al.* Fast, Scalable Generation of High-Quality Protein Multiple Sequence Alignments using Clustal Omega. *Molecular Systems Biology* **7**, 539 (2011).
228. Gasteiger, E. *et al.* Protein Identification and Analysis Tools on the Expasy Server. in *the Proteomics Protocols Handbook* (Ed. Walker, J. M.) 571–607 (Humana Press, 2005).
229. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic Local Alignment Search Tool. *Journal of Molecular Biology* **215**, 403–410 (1990).

230. Untergasser, A. *et al.* Primer3—New Capabilities and Interfaces. *Nucleic Acids Research* **40**, E115 (2012).
231. Owczarzy, R. *et al.* IDT Scitools: a Suite for Analysis and Design of Nucleic Acid Oligomers. *Nucleic Acids Research* **36**, W163–W169 (2008).
232. Nakamura, M. *et al.* Inhibitory Effects of Polyethers on Human Immunodeficiency Virus Replication. *Antimicrobial Agents and Chemotherapy* **36**, 492–494 (1992).
233. Takeuchi, T. *et al.* Compositions for Treating or Preventing Malaria and Method of Treating Malaria. (2005).
234. Blin, K. *et al.* AntiSMASH 4.0-Improvements in Chemistry Prediction and Gene Cluster Boundary Identification. *Nucleic Acids Research.* **45**, W36–W41 (2017).
235. Petković, H. *et al.* Substrate Specificity of the Acyl Transferase Domains of Epoc from the Epothilone Polyketide Synthase. *Organic and Biomolecular Chemistry* **6**, 500–506 (2008).
236. Wu, K., Chung, L., Revill, W. P., Katz, L. & Reeves, C. D. The FK520 Gene Cluster of *Streptomyces hygroscopicus* Var. Ascomyceticus (ATCC 14891) Contains Genes for Biosynthesis of Unusual Polyketide Extender Units. *Gene* **251**, 81–90 (2000).
237. Zhang, L. *et al.* Characterization of Giant Modular Pkss Provides Insight into Genetic Mechanism for Structural Diversification of Aminopolyol Polyketides. *Angewandte Chemie International Edition* **56**, 1740–1745 (2017).
238. Caffrey, P. Conserved Amino Acid Residues Correlating with Ketoreductase Stereospecificity in Modular Polyketide Synthases. *Chembiochem* **4**, 654–657 (2003).
239. Reid, R. *et al.* A Model of Structure and Catalysis for Ketoreductase Domains in Modular Polyketide Synthases. *Biochemistry* **42**, 72–79 (2003).
240. Du, Y. *et al.* Biosynthesis of the Apoptolidins in Nocardiopsis Sp. FU 40. *Tetrahedron* **67**, 6568–6575 (2011).
241. Haydock, S. F. *et al.* Organization of the Biosynthetic Gene Cluster for the Macrolide Concanamycin a in *Streptomyces neyagawaensis* ATCC 27449. *Microbiology* **151**, 3161–3169 (2005).
242. Rascher, A. *et al.* Cloning and Characterization of a Gene Cluster for Geldanamycin Production in *Streptomyces hygroscopicus* NRRL 3602. *FEMS Microbiology Letters* **218**, 223–230 (2003).
243. Rascher, A., Hu, Z., Buchanan, G. O., Reid, R. & Hutchinson, C. R. Insights into the Biosynthesis of the Benzoquinone Ansamycins Geldanamycin and Herbimycin,

Obtained by Gene Sequencing and Disruption. *Applied and Environmental Microbiology* **71**, 4862–4871 (2005).

244. Wenzel, S. C. *et al.* On the Biosynthetic Origin of Methoxymalonyl-Acyl Carrier Protein, the Substrate for Incorporation of "Glycolate" Units into Ansamitocin and Soraphen A. *Journal of the American Chemical Society* **128**, 14325–14336 (2006).
245. Sun, Y. *et al.* A Complete Gene Cluster from *Streptomyces nanchangensis* NS3226 Encoding Biosynthesis of the Polyether Ionophore Nanchangmycin. *Chemistry & Biology* **10**, 431–441 (2003).
246. Harvey, B. M. *et al.* Insights into Polyether Biosynthesis from Analysis of the Nigericin Biosynthetic Gene Cluster in *Streptomyces* Sp. DSM4137. *Chemistry & Biology* **14**, 703–714 (2007).
247. Cane, D. E., Liang, T.-C. & Hasler, H. Polyether Biosynthesis. Origin of the Oxygen Atoms of Monensin A. *Journal of the American Chemical Society* **103**, 5962–5965 (1981).
248. Sheldon, P. J., Busarow, S. B. & Hutchinson, C. R. Mapping the DNA-Binding Domain and Target Sequences of the *Streptomyces peucetius* Daunorubicin Biosynthesis Regulatory Protein, Dnri. *Molecular Microbiology* **44**, 449–460 (2002).
249. Wietzorrek, A. & Bibb, M. A Novel Family of Proteins that Regulates Antibiotic Production in Streptomycetes Appears to Contain an Ompr-Like DNA-Binding Fold. *Molecular Microbiology* **25**, 1181–1184 (1997).
250. Chen, J. & Xie, J. Role and Regulation of Bacterial LuxR-Like Regulators. *Journal of Cellular Biochemistry* **112**, 2694–2702 (2011).
251. Alekshun, M. N. & Levy, S. B. the Mar Regulon: Multiple Resistance to Antibiotics and Other Toxic Chemicals. *Trends in Microbiology* **7**, 410–413 (1999).
252. Wilkens, S. Structure and Mechanism of ABC Transporters. *F1000 Prime Reports* **7**, (2015).
253. Deochand, D. K. & Grove, A. Marr Family Transcription Factors: Dynamic Variations on a Common Scaffold. *Critical Reviews in Biochemistry and Molecular Biology* **52**, 595–613 (2017).
254. Murugan, E., Kong, R., Sun, H., Rao, F. & Liang, Z.-X. Expression, Purification and Characterization of the Acyl Carrier Protein Phosphodiesterase from *Pseudomonas Aeruginosa*. *Protein Expression and Purification* **71**, 132–138 (2010).

255. Thomas, J. & Cronan, J. E. The Enigmatic Acyl Carrier Protein Phosphodiesterase of *Escherichia coli* Genetic and Enzymological Characterization. *Journal of Biological Chemistry* **280**, 34675–34683 (2005).
256. Heathcote, M. L., Staunton, J. & Leadlay, P. F. Role of Type II Thioesterases: Evidence for Removal of Short Acyl Chains Produced by Aberrant Decarboxylation of Chain Extender Units. *Chemistry & Biology* **8**, 207–220 (2001).
257. Kotowska, M. & Pawlik, K. Roles of Type II Thioesterases and their Application for Secondary Metabolite Yield Improvement. *Applied Microbiology and Biotechnology* **98**, 7735–7746 (2014).
258. Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N. & Sternberg, M. J. E. the Phyre2 Web Portal for Protein Modeling, Prediction and Analysis. *Nature Protocols* **10**, 845–858 (2015).
259. Wang, P., Bashiri, G., Gao, X., Sawaya, M. R. & Tang, Y. Uncovering the Enzymes that Catalyze the Final Steps in Oxytetracycline Biosynthesis. *Journal of the American Chemical Society* **135**, 7138–7141 (2013).
260. Bosserman, M. A., Downey, T., Noinaj, N., Buchanan, S. K. & Rohr, J. Molecular Insight into Substrate Recognition and Catalysis of Baeyer–Villiger Monooxygenase Mtmoiv, the Key Frame-Modifying Enzyme in the Biosynthesis of Anticancer Agent Mithramycin. *ACS Chemical Biology* **8**, 2466–2477 (2013).
261. Kallio, P. *et al.* Tracing the Evolution of Angucyclinone Monooxygenases: Structural Determinants for C-12b Hydroxylation and Substrate Inhibition in Pgae. *Biochemistry* **52**, 4507–4516 (2013).
262. Frank, B. *et al.* Spiroketal Polyketide Formation in Sorangium: Identification and Analysis of the Biosynthetic Gene Cluster for the Highly Cytotoxic Spirangienes. *Chemistry & Biology* **14**, 221–233 (2007).
263. Gaitatzis, N. *et al.* The Biosynthesis of the Aromatic Myxobacterial Electron Transport Inhibitor Stigmatellin is Directed by a Novel Type of Modular Polyketide Synthase. *Journal of Biological Chemistry* **277**, 13082–13090 (2002).
264. Perlova, O., Gerth, K., Kaiser, O., Hans, A. & Müller, R. Identification and Analysis of the Chivosazol Biosynthetic Gene Cluster from the Myxobacterial Model Strain *Sorangium Cellulosum* So Ce56. *Journal of Biotechnology* **121**, 174–191 (2006).
265. Tang, L. *et al.* Elucidating the Mechanism of *cis* Double Bond Formation in Epothilone Biosynthesis. *Journal of the American Chemical Society* **126**, 46–47 (2004).

266. Xu, W. *et al.* An Iterative Module in the Azalomycin F Polyketide Synthase Contains a Switchable Enoylreductase Domain. *Angewandte Chemie International Edition* **56**, 5503–5506 (2017).
267. Siskos, A. P. *et al.* Molecular Basis of Celmer's Rules: Stereochemistry of Catalysis by Isolated Ketoreductase Domains from Modular Polyketide Synthases. *Chemistry & Biology* **12**, 1145–1153 (2005).
268. Fonseca, T., Lopes, N. P., Gates, P. J. & Staunton, J. Fragmentation Studies on Tetronasin by Accurate-Mass Electrospray Tandem Mass Spectrometry. *Journal of the American Society for Mass Spectrometry* **15**, 325–335 (2004).
269. Bierman, M. *et al.* Plasmid Cloning Vectors for the Conjugal Transfer of DNA from *Escherichia coli* to *Streptomyces* Spp. *Gene* **116**, 43–49 (1992).
270. Huijbers, M. M. E., Montersino, S., Westphal, A. H., Tischler, D. & Van Berkel, W. J. H. Flavin Dependent Monooxygenases. *Archives of Biochemistry and Biophysics* **544**, 2–17 (2014).
271. Ortiz De Montellano, P. R. Hydrocarbon Hydroxylation by Cytochrome P450 Enzymes. *Chemical Reviews* **110**, 932–948 (2010).
272. Myronovskyi, M. & Luzhetskyy, A. Native and Engineered Promoters in Natural Product Discovery. *Natural Product Reports* **33**, 1006–1019 (2016).
273. Bilyk, B., Horbal, L. & Luzhetskyy, A. Chromosomal Position Effect Influences the Heterologous Expression of Genes and Biosynthetic Gene Clusters in *Streptomyces albus* J1074. *Microbial Cell Factories* **16**, 5 (2017).
274. Gelin, S. & Pollet, P. Tautomerism in Acyl Tetronic Acids. *Tetrahedron Letters* **21**, 4 (1980).
275. Johnson, J. D., Creighton, D. J. & Lambert, M. R. Stereochemistry and Function of Oxaloacetate Keto-Enol Tautomerase. *Journal of Biological Chemistry* **261**, 4535–4541 (1986).
276. Rétey, J., Bartl, K., Ripp, E. & Hull, W. E. Stereospecificity of Phenylpyruvate Tautomerase. A Convenient Method for the Preparation of Chirally Labelled Phenylpyruvates. *European Journal of Biochemistry* **72**, 251–257 (1977).
277. Widlanski, T., Bender, S. L. & Knowles, J. R. Dehydroquinate Synthase: A Sheep in Wolf's Clothing? *Journal of the American Chemical Society* **111**, 2299–2300 (1989).
278. Yoeun, S., Cho, K. & Han, O. Structural Evidence for the Substrate Channeling of Rice Allene Oxide Cyclase in Biologically Analogous Nazarov Reaction. *Frontiers in Chemistry* **6**, (2018).

279. Bogart, J. W. & Bowers, A. A. Thiopeptide Pyridine Synthase Tbtd Catalyzes an Intermolecular Formal Aza-Diels-Alder Reaction. *Journal of the American Chemical Society* (2019).
280. Ohashi, M. *et al.* SAM-Dependent Enzyme-Catalysed Pericyclic Reactions in Natural Product Biosynthesis. *Nature* **549**, 502–506 (2017).
281. Parascandolo, J. S. *et al.* Insights into 6-Methylsalicylic Acid Bio-Assembly by using Chemical Probes. *Angewandte Chemie International Edition* **55**, 3463–3467 (2016).
282. Tosin, M. *et al. in vivo* Trapping of Polyketide Intermediates from an Assembly Line Synthase using Malonyl Carba(Dethia)-N-Acetyl Cysteamines. *Chemical Communications* **47**, 3460–3462 (2011).
283. Wilkening, I. *et al.* Second-Generation Probes for Biosynthetic Intermediate Capture: Towards a Comprehensive Profiling of Polyketide Assembly. *Chemical Communications* (2016).
284. Pang, B., Wang, M. & Liu, W. Cyclization of Polyketides and Non-Ribosomal Peptides on and off their Assembly Lines. *Natural Product Reports* (2015).
285. Zhang, Z. *et al.* Enzyme-Catalyzed Inverse-Electron Demand Diels–Alder Reaction in the Biosynthesis of Antifungal Ilicicolin H. *Journal of the American Chemical Society* (2019).
286. Aharoni, A. *et al.* The 'Evolvability' of Promiscuous Protein Functions. *Nature Genetics* **37**, 73–76 (2005).
287. Kato, N. *et al.* A New Enzyme Involved in the Control of the Stereochemistry in the Decalin Formation During Equisetin Biosynthesis. *Biochemical and Biophysical Research Communications* **460**, 210–215 (2015).
288. Kato, N. *et al.* Control of the Stereochemical Course of [4+2] Cycloaddition During *trans*-Decalin Formation by Fsa2-Family Enzymes. *Angewandte Chemie International Edition Engl.* **57**, 9754–9758 (2018).
289. Less, S. L., Handa, S., Leadlay, P. F., Dutton, C. J. & Staunton, J. Biosynthesis of Tetronasin: Part 5. Novel Fluorinated and Non-Fluorinated Analogues of Tetronasin via Intact Incorporation of Di-, Tri-and Tetraketide Analogue Precursors. *Tetrahedron Letters* **37**, 3511–3514 (1996).
290. Luhavaya, H. Biosynthesis of Salinomycin, A Polyether Ionophore with Anticancer and Antituberculosis Activities. (University of Cambridge, 2015).

291. Wlodawer, A., Minor, W., Dauter, Z. & Jaskolski, M. Protein Crystallography for Aspiring Crystallographers or how to Avoid Pitfalls and Traps in Macromolecular Structure Determination. *The FEBS Journal* **280**, 5705–5736 (2013).
292. Russo Krauss, I., Merlino, A., Vergara, A. & Sica, F. An Overview of Biological Macromolecule Crystallization. *International Journal of Molecular Sciences* **14**, 11643–11691 (2013).
293. Costa, K. C., Glasser, N. R., Conway, S. J. & Newman, D. K. Pyocyanin Degradation by a Tautomerizing Demethylase Inhibits *Pseudomonas aeruginosa* Biofilms. *Science* **355**, 170–173 (2017).
294. Lebowitz, J., Lewis, M. S. & Schuck, P. Modern Analytical Ultracentrifugation in Protein Science: a Tutorial Review. *Protein Science* **11**, 2067–2079 (2002).
295. Cowtan, K. Phase Problem in X-Ray Crystallography, and its Solution. in *Encyclopedia of Life Sciences* (Ed. John Wiley & Sons, Ltd) (John Wiley & Sons, Ltd, 2003).
296. Neumann, P. *et al.* Crystal Structures of Physcomitrella Patens AOC1 and AOC2: Insights into the Enzyme Mechanism and Differences in Substrate Specificity. *Plant Physiol* **160**, 1251–1266 (2012).
297. Taylor, G. L. Introduction to Phasing. *Acta Crystallographica. Section D, Biological Crystallography* **66**, 325–338 (2010).
298. Wingfield, P. *N*-Terminal Methionine Processing. *Current Protocols in Protein Science* **88**, 6.14.1-6.14.3 (2017).
299. Laskowski, R. A. *et al.* Pdbsum: A Web-Based Database of Summaries and Analyses of All PDB Structures. *Trends in Biochemical Sciences* **22**, 488–490 (1997).
300. Hotta, K. *et al.* Enzymatic Catalysis of Anti-Baldwin Ring Closure in Polyether Biosynthesis. *Nature* **483**, 355–358 (2012).
301. Wang, L. *et al.* A Rieske Oxygenase/Epoxide Hydrolase-Catalysed Reaction Cascade Creates Oxygen Heterocycles in Mupirocin Biosynthesis. *Nature Catalysis* **1**, 968 (2018).
302. Taguchi, T. *et al.* Remarkably Different Structures and Reaction Mechanisms of Ketoreductases for the Opposite Stereochemical Control in the Biosynthesis of BIQ Antibiotics. *Bioorganic & Medicinal Chemistry* **12**, 5917–5927 (2004).
303. Greenfield, N. J. using Circular Dichromism Spectra to Estimate Protein Secondary Structure. *Nature Protocols* **1**, 2876–2890 (2006).

304. Chang, Z. *et al.* Crystal Structure of LepI, a Multifunctional SAM-Dependent Enzyme Which Catalyzes Pericyclic Reactions in Leporin Biosynthesis. *Organic & Biomolecular Chemistry* **17**, 2070–2076 (2019).
305. Chang, M. *et al.* Crystal Structure of the Multifunctional SAM-Dependent Enzyme LepI Provides Insights into its Catalytic Mechanism. *Biochemical and Biophysical Research Communications* (2019) [electronic prepublication].
306. Qiu, X. *et al.* Crystal Structure of Beta-Ketoacyl-Acyl Carrier Protein Synthase III. A Key Condensing Enzyme in Bacterial Fatty Acid Biosynthesis. *Journal of Biological Chemistry* **274**, 36465–36471 (1999).
307. Davies, C., Heath, R. J., White, S. W. & Rock, C. O. The 1.8 Å Crystal Structure and Active-Site Architecture of β-Ketoacyl-Acyl Carrier Protein Synthase III (Fabh) from *Escherichia coli*. *Structure* **8**, 185–195 (2000).

Appendix

Appendix Notes

Appendix note 1

The sequence of Tsn15. Prior to this thesis, a different gene in the tetronasin gene cluster, an ABC transporter, was named Tsn15. The Tsn15 described in this paper was not annotated. As such, the ABC transporter has been renamed Tsn15b and Tsn15 now refers to the [4+2] cyclase homologue. For clarity the DNA and protein sequence of Tsn15 are presented below:

>tsn15

ATGACCACTTCCATCGATCCGACGACCCCGCTGACCTACAACCCCGTCATCGACGCGCTCGTCGGCTCGTGGCGCCAGAT
CATCGACGCCGACTACTCGGCGGACGACACCCGGCTGCCCGATCTCGCCGTGCTGGCCCGCTCCACCGCGCGGGCGGTCG
CGGCTGCCGTACCCCGTCCGCTCGCGGAGATCTCGGCCCCGGACGCGCCGGACGAGCGCGGCGAACTCGTGCTGCTGGAG
AAGGTGATCCAGGAAGTGGCCGACCGCGAGTACACCCCGCTGAGCCCCGAGGGGCCGAGCGTCGGGGACCTCGTCCTGGT
GACGGAGAAGATCTACAACTCCGACCGCGAGGAGATCGGCGCGGACACCGGGCGGCTGCGGATCATCCGCAAGGACCCGG
AGACCGGGCACCACTTCACGGTCTCGCTCGTCACGTCCACGGTGCAGGGCAACAAGCTGTTCGCCTTCGGCTACACCGAG
ATGGAGGCCCAGCTCGCCGGGGGCCGCACCACCATCCAGGTCGCCTGCTGGGACGGCCCCTGGGCCGGCATGAGCGGCAC
CCTGTCCTGGGTCATCAACTCCATGACGGCCGCCGAGTCGCGGTACGAGCTGCGCCGCTGA

>Tsn15

MTTSIDPTTPLTYNPVIDALVGSWRQIIDADYSADDTRLPDLAVLARSTARAVAAAVPRPLAEISAPDAPDERGELVLLE
KVIQEVADREYTPLSPEGPSVGDLVLVTEKIYNSDREEIGADTGRLRIIRKDPETGHHFTVSLVTSTVQGNKLFAFGYTE
MEAQLAGGRTTIQVACWDGPWAGMSGTLSWVINSMTAAESRYELRR

Appendix note 2

Revised sequence of Tsn11. Closer examination of the *tsn11* gene compared to the deposited sequence revealed the start codon to be an additional 12 bp 5' of the previously annotated start codon. The new start codon was a methionine (ATG) rather than a leucine (TTG) and was six bp downstream from a likely Shine-Delgarno sequence (GGAAGAA). The revised protein sequence of the updated Tsn11 is presented below.

>Tsn11

MEIPLTGTVVIAGAGPVGLFLASELRLAGVEAVVLERSPKANEHTVGGTLHARTADLFDQRGIMDTLRAGNPPLWPRLHF
ASYWLDLAPHMEDEYSLLLPQQYTEEMLEAHATELGADIRRGHTCVSLTQDADGVTVGVRADSGDYELRGAYLVGCDGGD
STVRELAAFPVQESGPRWYGLLADVESIEGDWHPGNYPGGQFAVIRSPHEGGPSRIMTLEFNETTQPPPADQPVSVEEVI
ASTERITGRTPVVGEVQWLHRYTNTTREAENYRQGRVFVAGDAAHLHVAFAGHGLSTGLHDAANLGWKLAAVLDGRAPDS
LLDTYDEERRPVGHRACVFTQSQMALLTQGQQLDILRQLFTDLVKLPEVNHHLITTVTDVRYALDGAEKEDTHPLLGRPV
PNQLVKDADGQATAVAEALRAGRGVLIDLTDGAAALPDTSGRRGHLDSVSRGPADAVDATALLVRPDGFVAWAATADTGN
DGLEPALRRWFGDTA

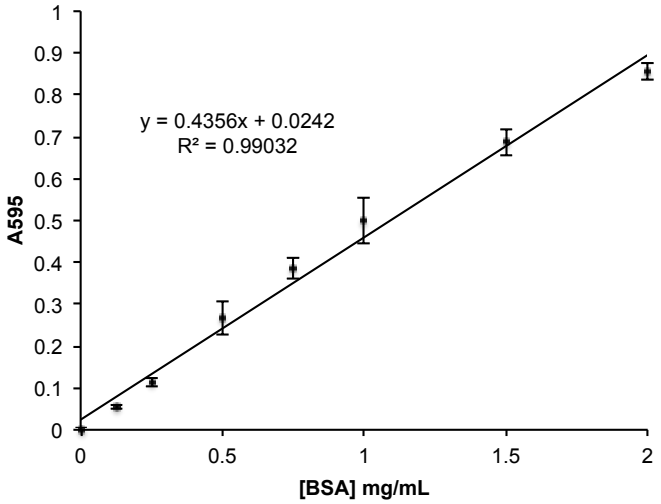
Appendix note 3

Revised sequence of Tmn9. During cloning of *tmn9* it was noted that the sequence differed slightly from that previously published[15] (confirmed by sequencing independent PCR reactions). The revised Tmn9 protein sequence is as follows:

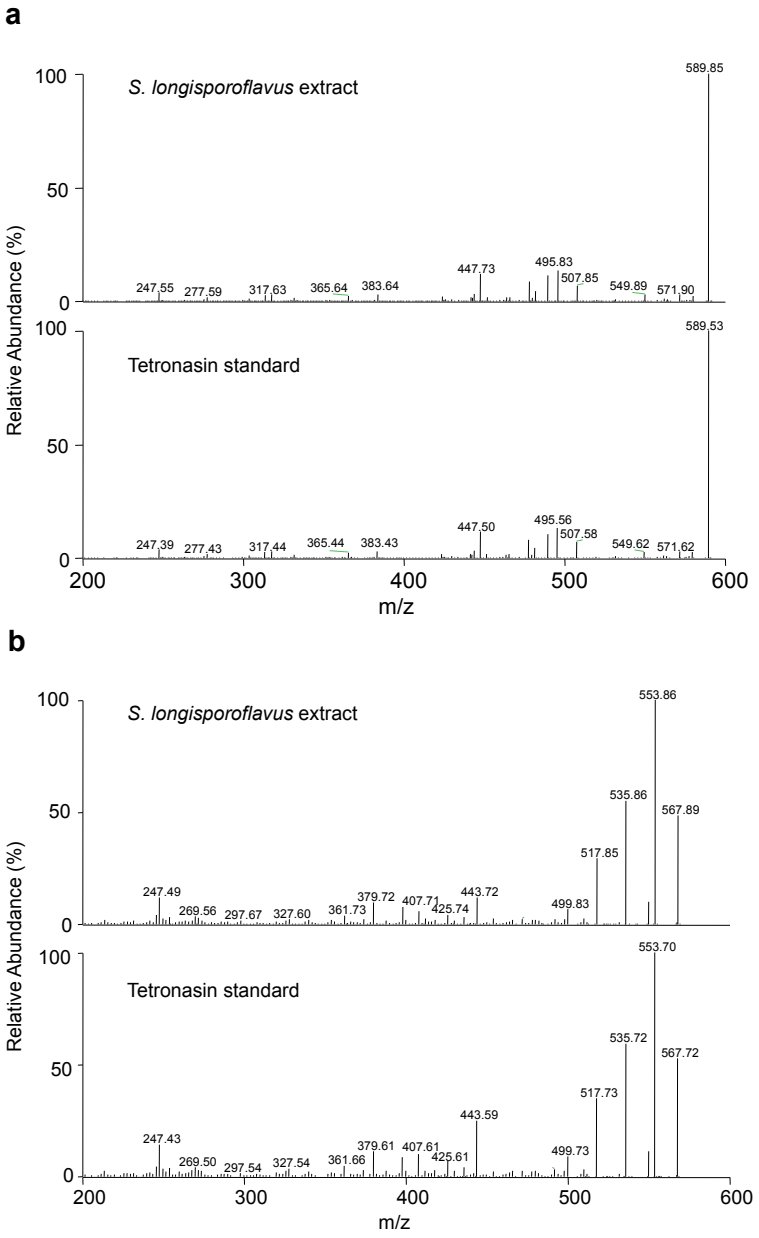
>Tmn9

MSEPVVVVGAGPAGLMLACELAMRDVPAVLVDIHPTQRAEAPAMAINAGTLEMLDQRGLAAGLREGTVTFPEVRFADLRL
AFEKVQGPREPTHMVLQSRLEKVLIDRAVELGVDLRWATRLTGFEEAADGSGVTVTLASDAGEEQLRCRYLVGCDGRESI
VRKQAGIDYVGDDWVIVRGIVGDVAINREDVAPEQYGLSYTDNGDQFLGAPLSPDVMRVFSAEFSTEPPEFEDGPATLEQ
LGDAVKRLTGKELKATEAHWLQHYSIVTRNAEQYRKGRVFIAGDAAHVHYPYNGQGLGTAIGDAVNLGWKIAAEVHGWAP
ADLLDSYHVERHLAGRLACMNIQAQLALLYPRPLARYMREMMGEFLKFDEVNVFLAEIVTNLGPAVPIAYEGVPEPVEGD
RLLGRRLPKVQIKTADGDMGVAETLQSGRGVLLDLSGDASAQEESGWADRVDVVRAQPVPDLPGTLLLRPDGCVAWHDGG
GWGQDELRTALRTWFGAPTG

Appendix Figures

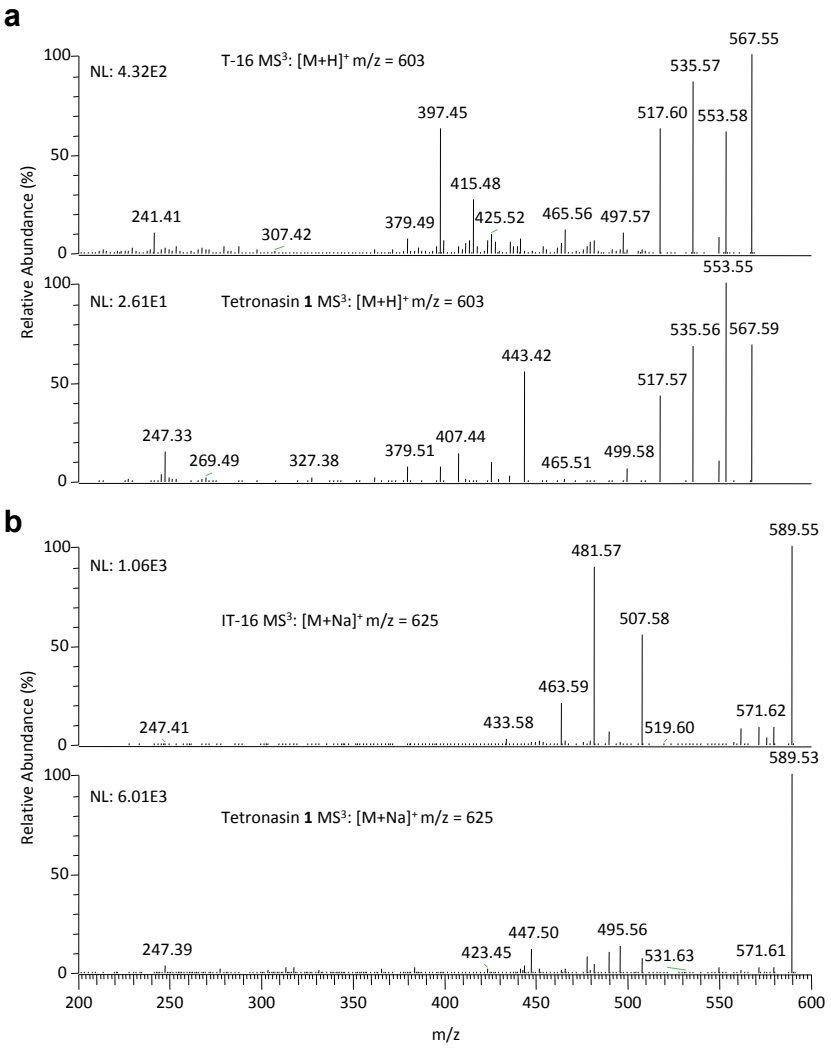


Appendix Figure 1
BSA standard curve using Bradford reagent. A BSA standard curve was prepared using the Bradford reagent. n = 3. Error bars represent the standard deviation.

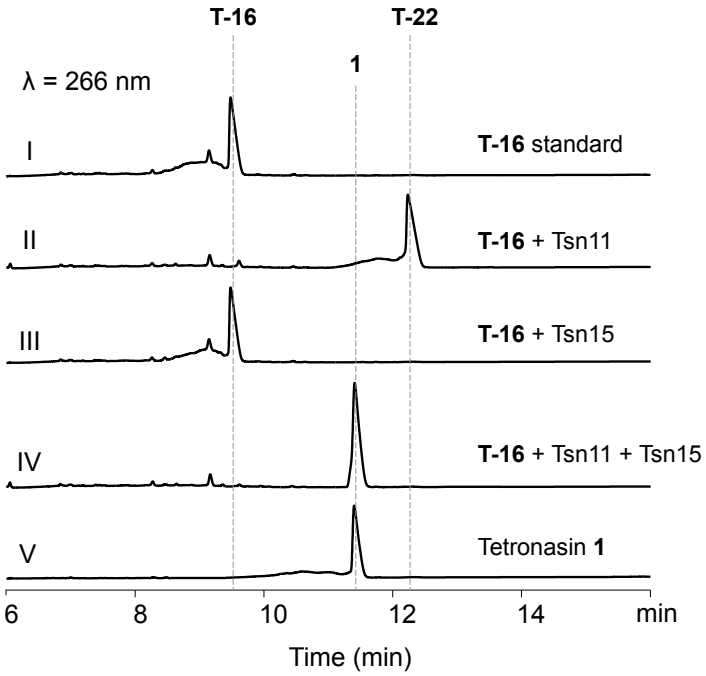


Appendix Figure 2

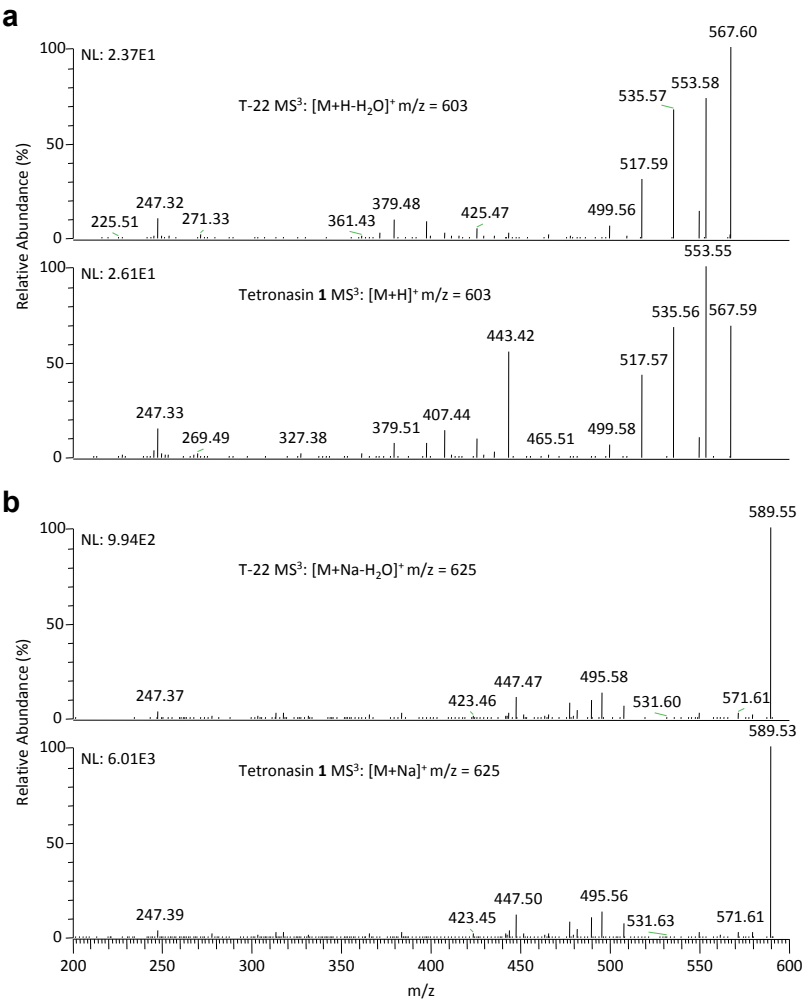
MS^3 fragmentation of tetronasin produced by *S. longisporoflavus.* The MS^3 fragmentation patterns of the $[M+Na]^+$ and the $[M+H]^+$ ions of the tetronasin produced by *S. longisporoflavus* were compared to the same MS^3 fragments from a tetronasin standard. **a,** The MS^3 spectra from the $[M+Na]^+$ ion (625 -> 607 ($-H_2O$) -> fragments) for tetronasin produced in vivo by *S. longisporoflavus* (top) and a standard (bottom). **b,** The MS^3 spectra from the $[M+H]^+$ ion (603 -> 585 ($-H_2O$) -> fragments) for tetronasin produced in vivo by *S. longisporoflavus* (top) and a standard (bottom). The detected fragments were in good accordance with those previously observed for tetronasin[268].



Appendix Figure 3
MS^3 mass spectra of intermediate T-16. a, Top: MS^3 spectra of the T-16 m/z = 603 $[M+H-H_2O]^+$ ion: m/z = 603 -> 585 ($-H_2O$) -> fragments. Bottom: MS^3 spectra of the tetronasin **1** m/z = 603 $[M+H]^+$ ion: m/z = 603 -> 585 ($-H_2O$) -> fragments. **b,** Top: MS^3 spectra of the T-16 m/z = 625 $[M+Na-H_2O]^+$ ion: m/z = 625 -> 607 ($-H_2O$) -> fragments. Bottom: MS^3 spectra of the tetronasin 1 m/z = 603 $[M+Na]^+$ ion: m/z = 625 -> 607 ($-H_2O$) -> fragments.

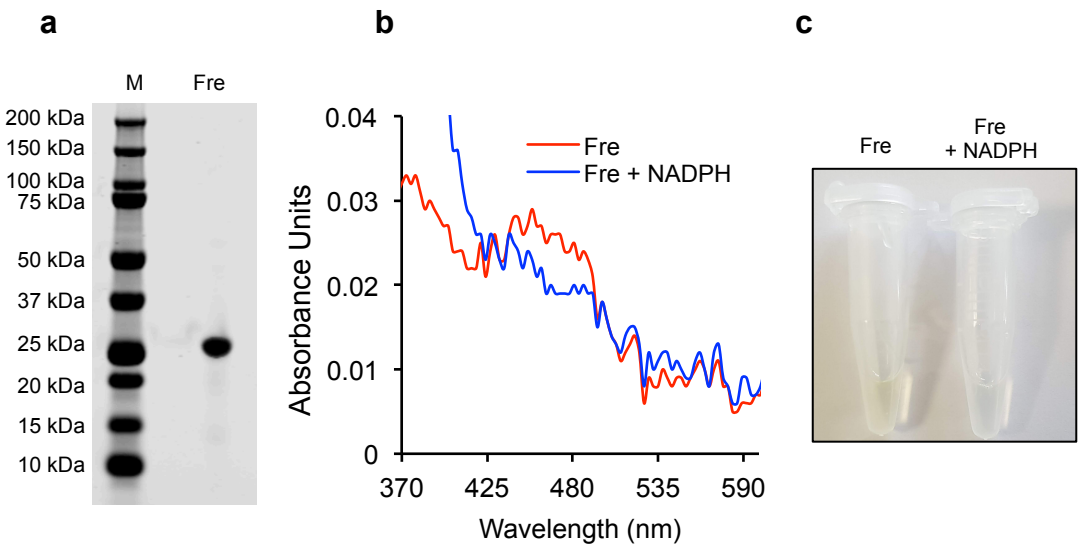


Appendix Figure 4
In vitro conversion of T-16 into tetronasin using purified Tsn11 and Tsn15. The *in vitro* conversion of T-16 into tetronasin **1** by Tsn11 and Tsn15 was monitored by HPLC analysis. (I) T-16 incubated in the absence of enzymes. (II) Conversion of T-16 into T-22 using Tsn11 (III) No reaction occurred when T-16 was incubated with Tsn15 alone. (IV) Conversion of T-16 into tetronasin **1** using Tsn11 and Tsn15. (V) Conversion of T-22 into tetronasin **1** using Tsn15. (VI) Tetronasin **1** standard. Data are representative of three independent experiments.

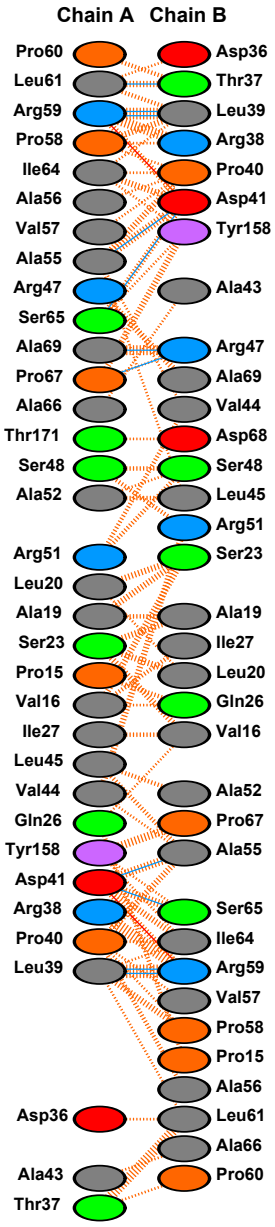


Appendix Figure 5

MS^3 spectra of the T-22. **a,** Top: MS^3 spectra of the T-22 m/z = 603 $[M+H-H_2O]^+$ ion: m/z = 603 -> 585 ($-H_2O$) -> fragments. Bottom: MS^3 spectra of the tetronasin **1** m/z = 603 $[M+H]^+$ ion: m/z = 603 -> 585 ($-H_2O$) -> fragments. **b,** Top: MS^3 spectra of the T-22 m/z = 625 $[M+Na-H_2O]^+$ ion: m/z = 625 -> 607 ($-H_2O$) -> fragments. Bottom: MS^3 spectra of the tetronasin **1** m/z = 603 $[M+Na]^+$ ion: m/z = 625 -> 607 ($-H_2O$) -> fragments.

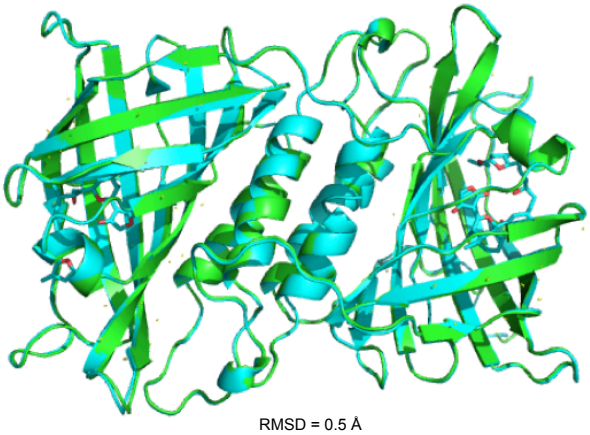


Appendix Figure 6
Purification and activity of Fre. **a,** Fre cloned into pET28a(+) and purified from *E.coli* using Ni-affinity. **b,** Fre co-purified with a flavin cofactor bound to it. Adding NADPH resulted in a detectable drop in UV/Vis absorbance and decolourised the protein, indicating Fre was active **c,** Decolouration of Fre upon adding NADPH.

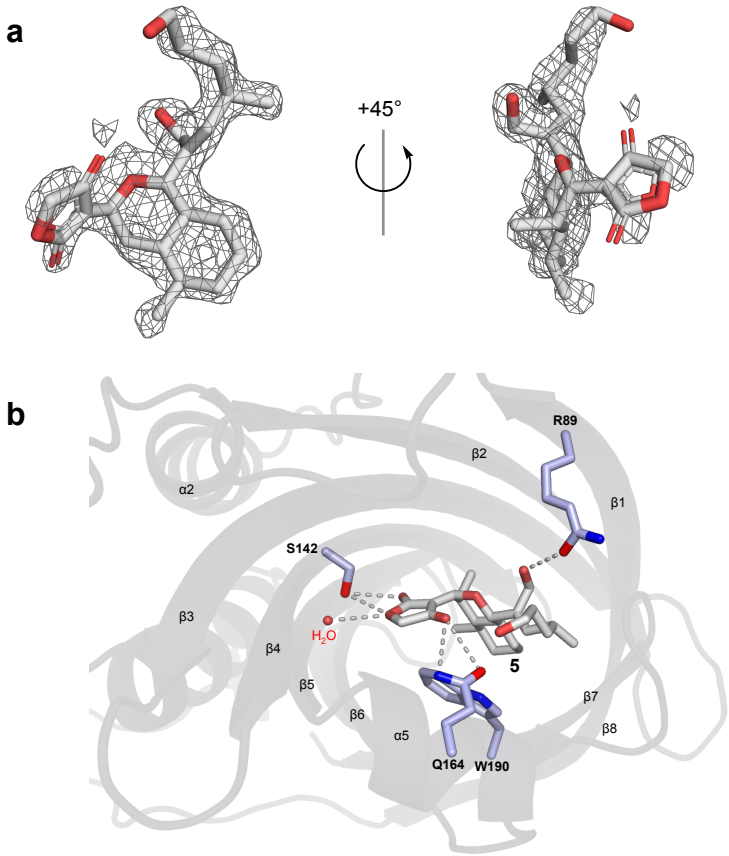


Appendix Figure 7

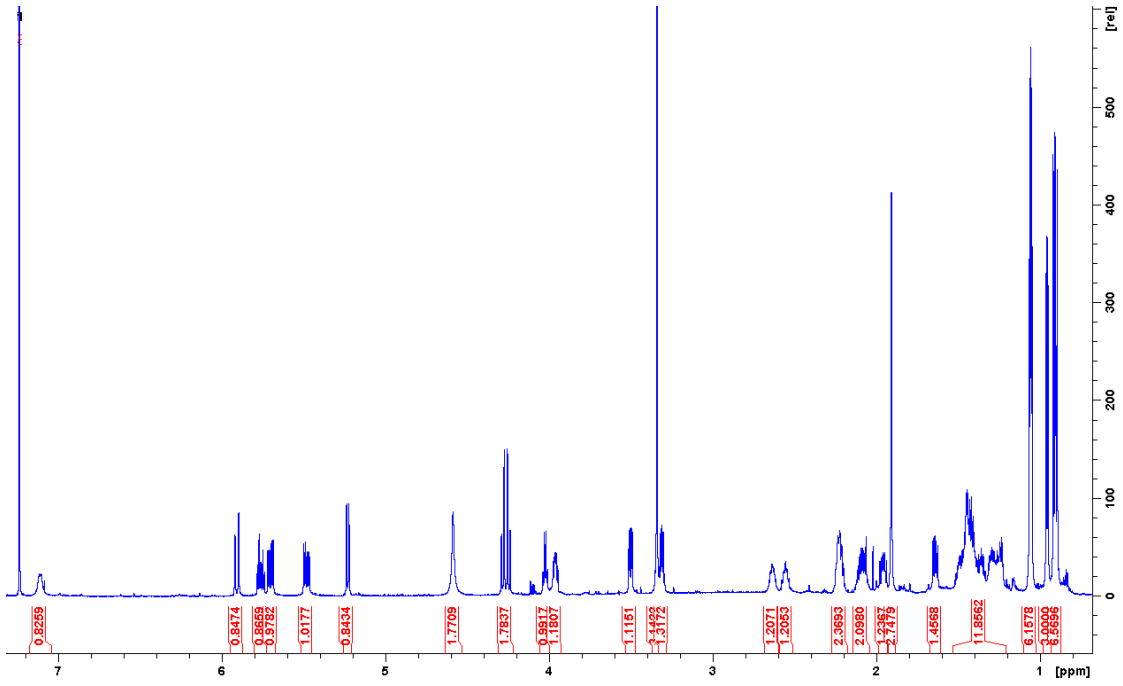
Dimerisation interface of Tsn15 (Figure made by Fernanda Paiva). The online structural biology tool PDBsum[299] (EMBL-EBI) was used to generate a diagram showing the various interactions between the two chains of Tsn15. Orange dashes = non-polar interactions. The thickness of the dashes is proportional to the number of atomic interactions. Blue line = hydrogen bonds. Red lines = salt bridges. Colour of ellipses: red = negatively charged amino acid; blue = positively charged amino acid; magenta = aromatic amino acid; green = polar amino acid; grey = non-polar amino acid.



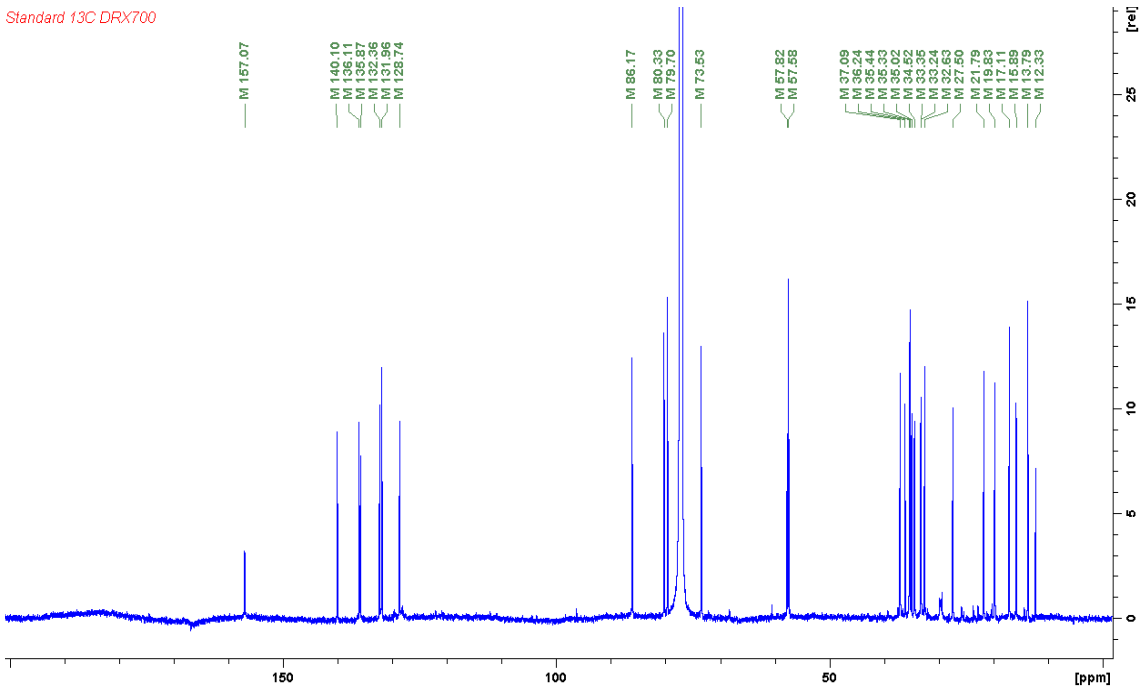
Appendix Figure 8
Comparison of the Cα chains between Tsn15 and Tsn15-ligand complex. (Figure made by Fernanda Paiva). Aligning the two structures revealed no significant changes to the Cα positions upon substrate binding. Green = Tsn15. Cyan = Tsn15-ligand complex.



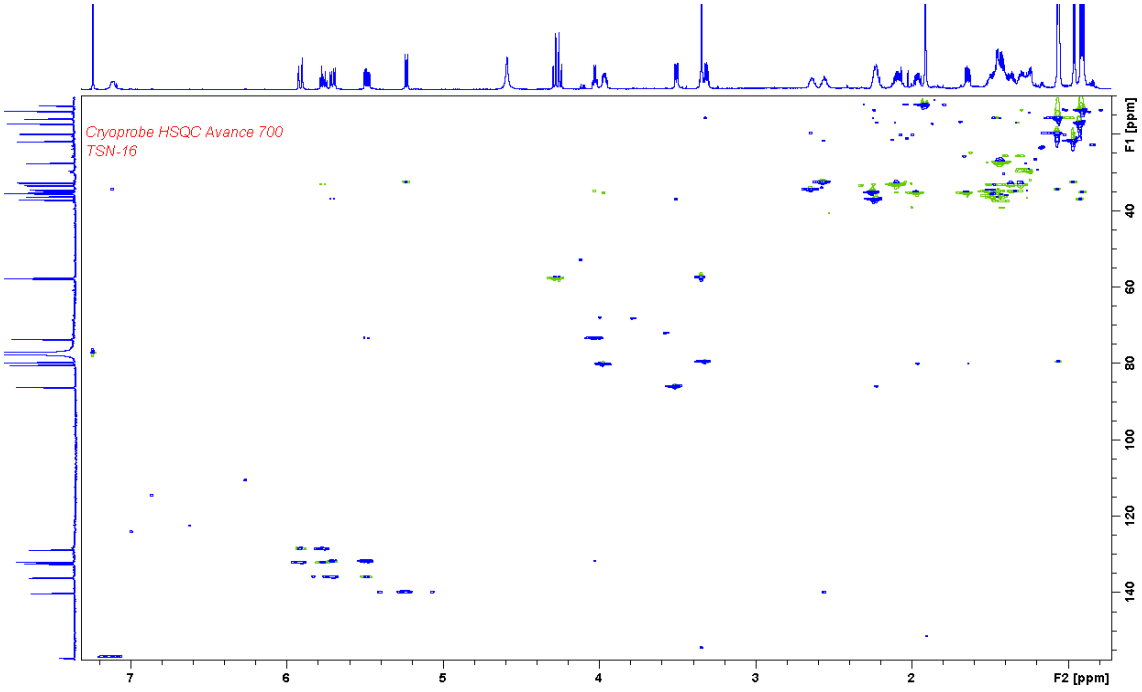
Appendix Figure 9
Chain B of the Tsn15-ligand complex. **a,** Resolvable portion of dT-22 present in chain B of Tsn15 (F_o-F_c contoured to 2.0 σ). The electron density of the tetrahydrofuran portion of the molecule (from C17 onwards) was too poor to be accurately modelled. **b,** Ribbon representation of chain B of the Tsn15-ligand complex.



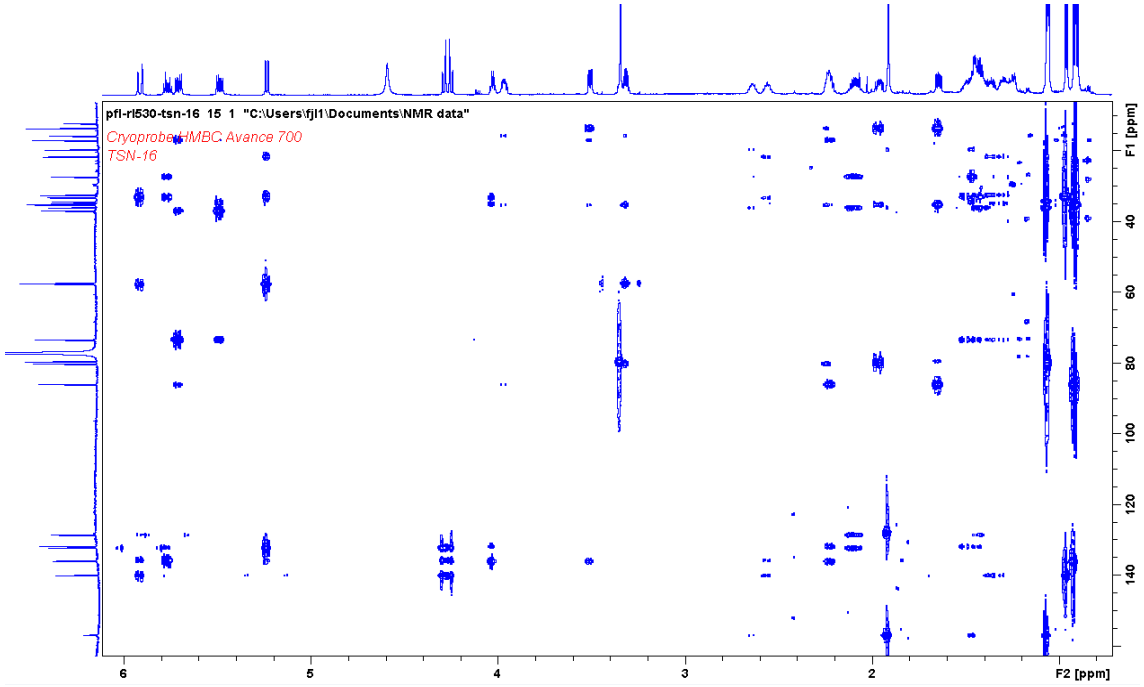
Appendix Figure 10
^{1}H NMR spectrum of T-16 in $CDCl_3$.



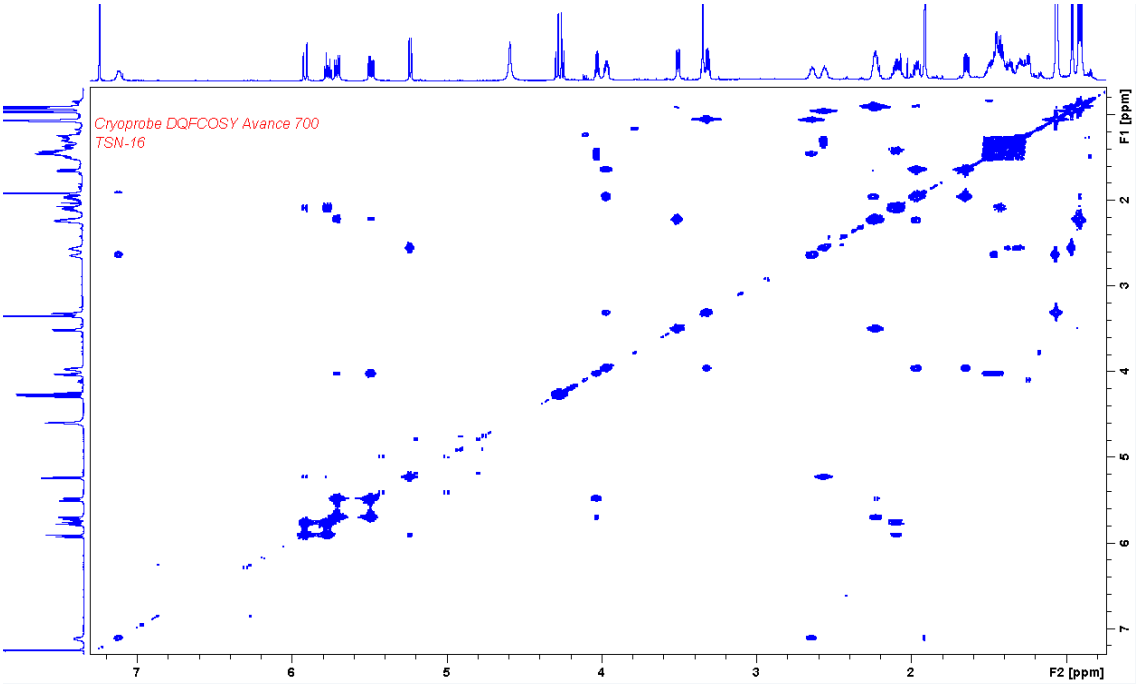
Appendix Figure 11
^{13}C NMR spectrum of T-16 in $CDCl_3$.



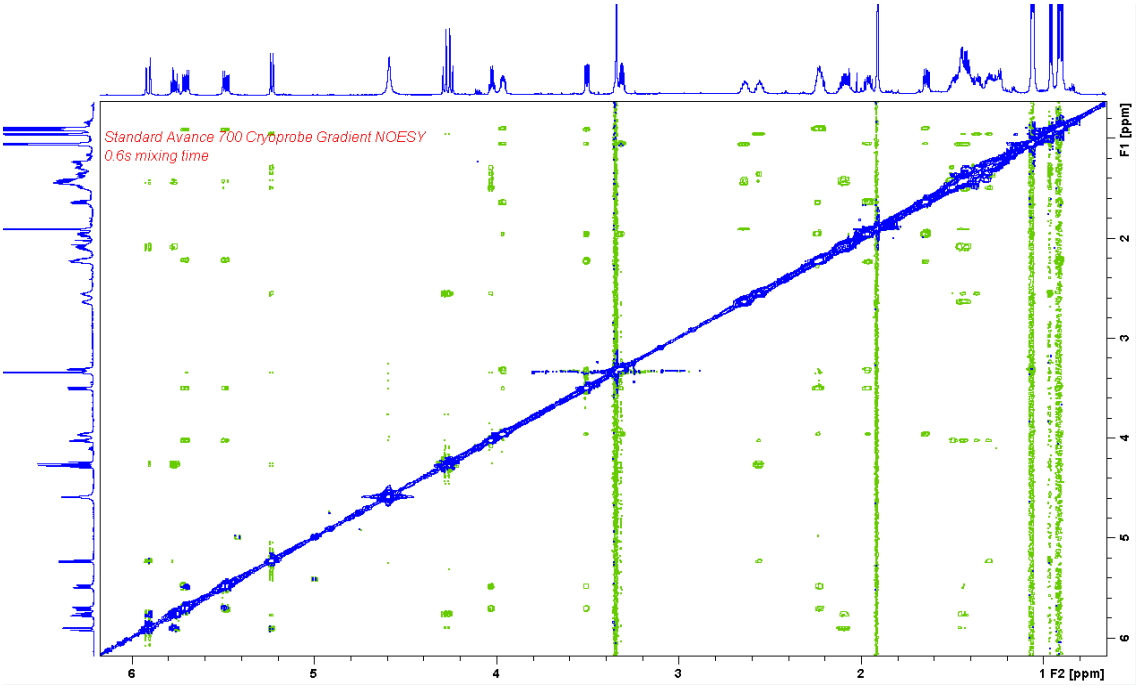
Appendix Figure 12
HSQC spectrum of T-16 in $CDCl_3$. Green peaks are CH_2 groups, and blue peaks are CH or CH_3.



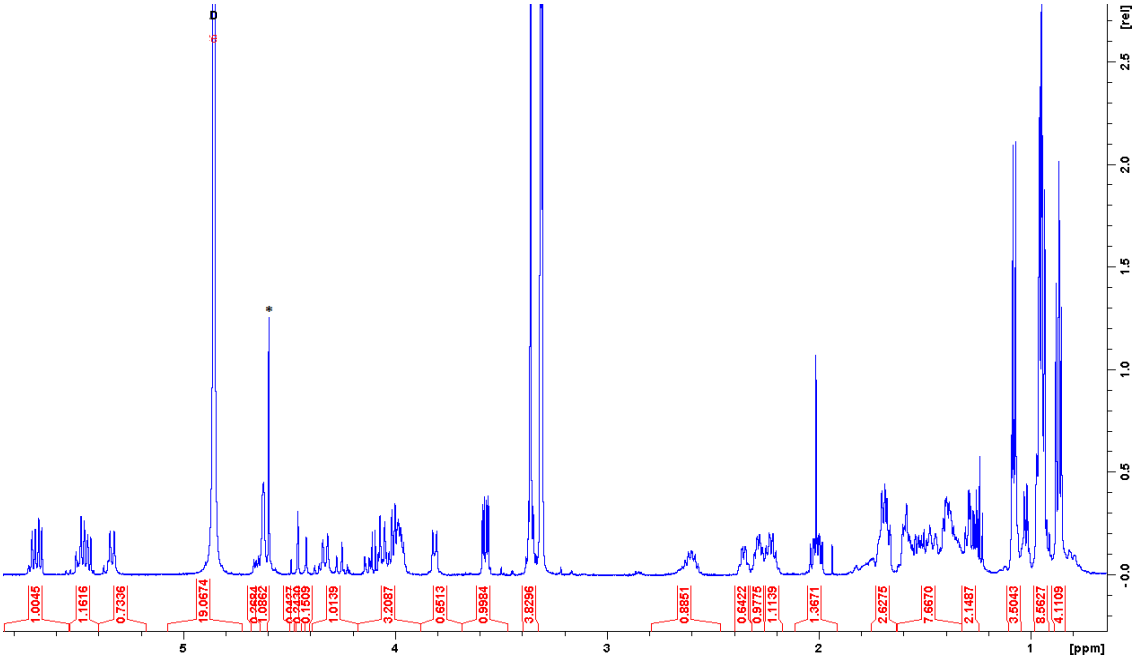
Appendix Figure 13
HMBC spectrum of T-16 in $CDCl_3$.



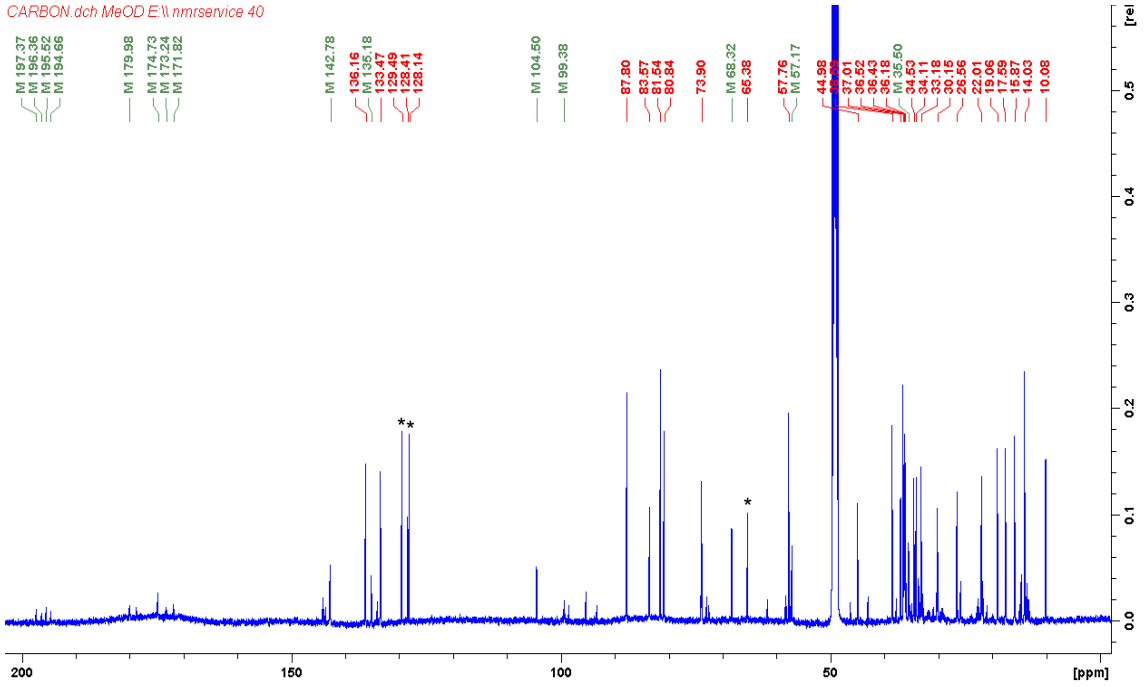
Appendix Figure 14
^{1}H-^{1}H COSY spectrum of T-16 in in $CDCl_3$.



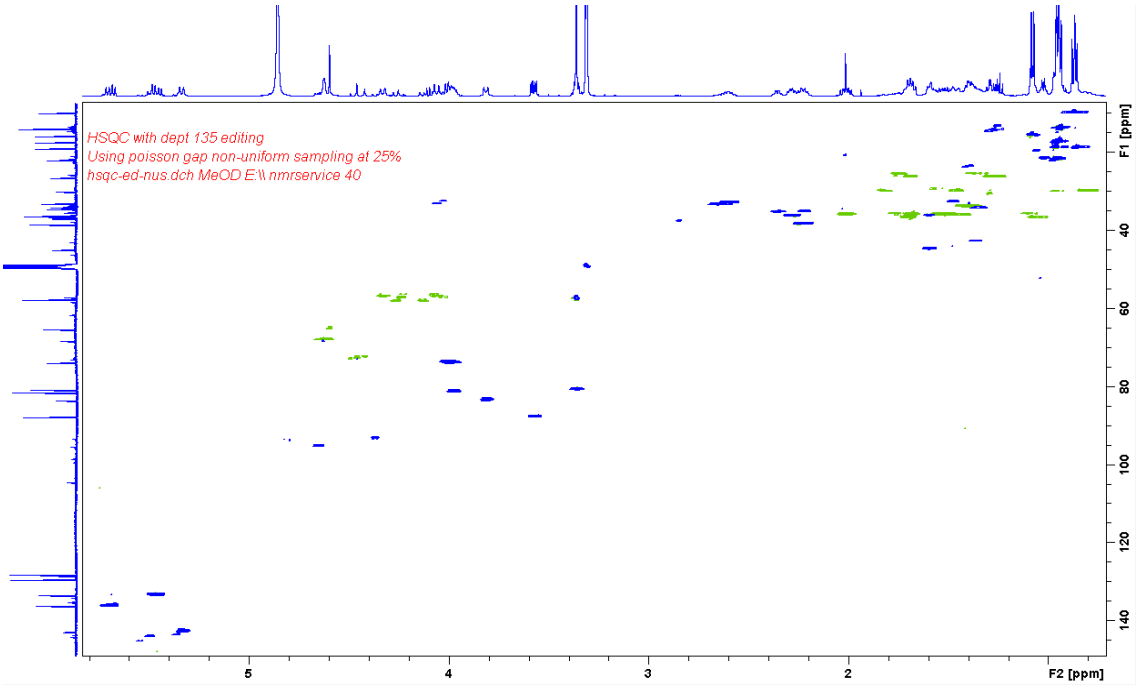
Appendix Figure 15
^{1}H-^{1}H NOESY spectrum of T-16 in $CDCl_3$.



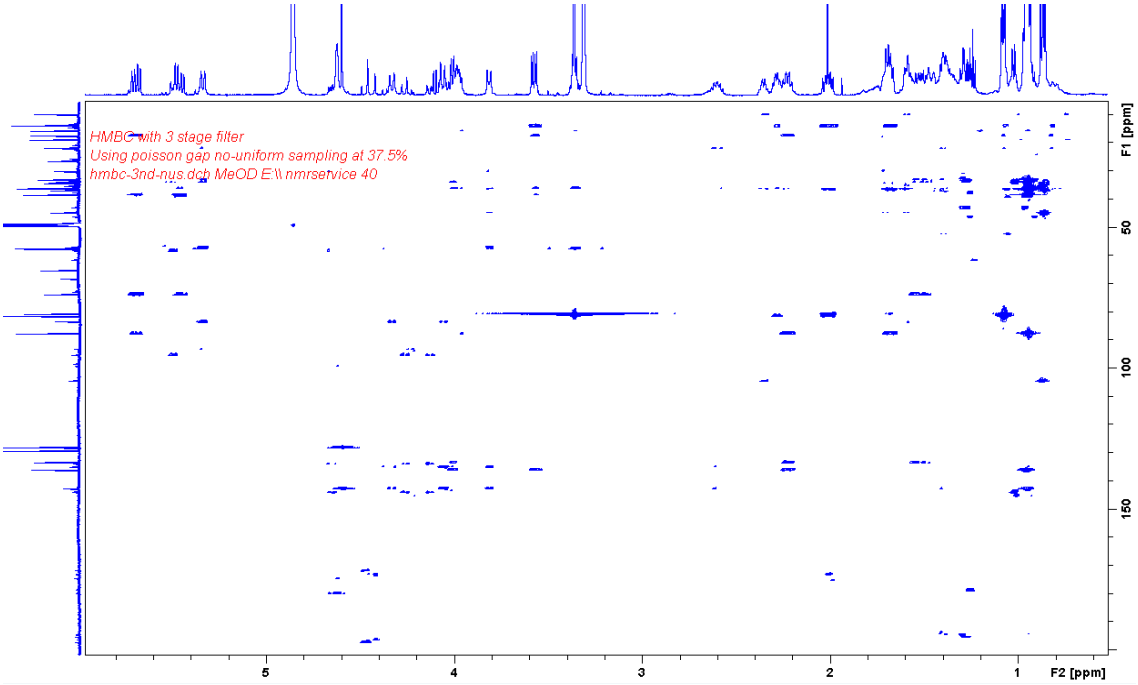
Appendix Figure 16
^{1}H NMR spectrum of T-22 in CD_3OD. *The peak at d 4.60 is due to a minor contaminant $PhCH_2OH$ (aromatic peaks at d 7.2-7.4 ppm not shown).



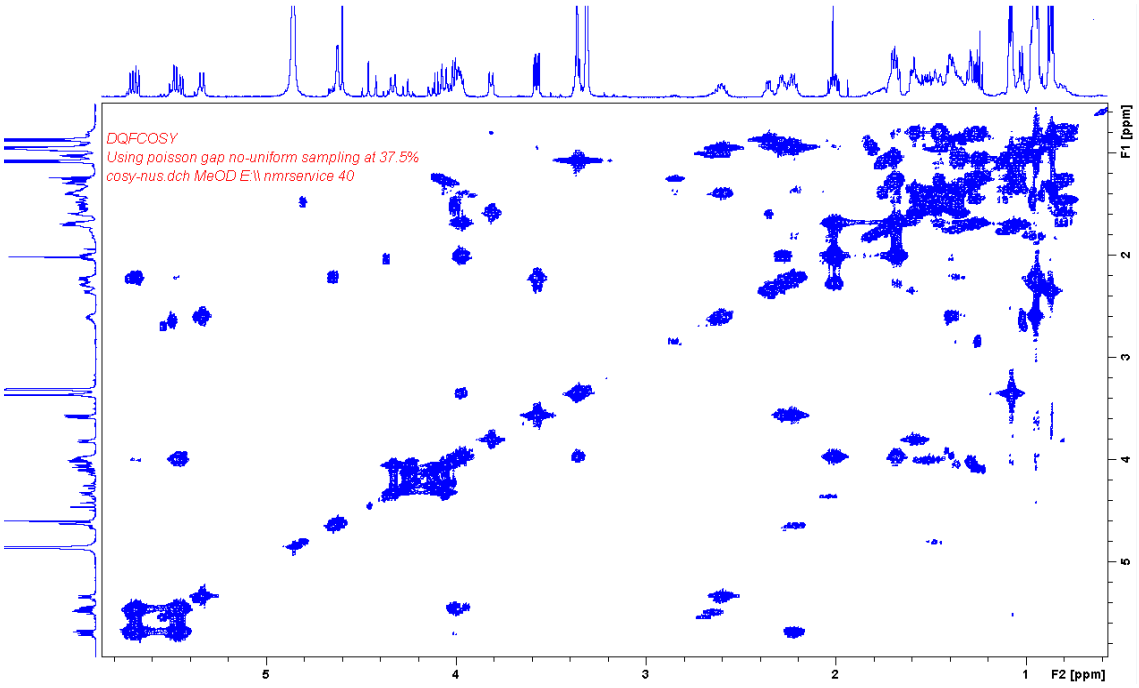
Appendix Figure 17
^{13}C NMR spectrum of T-22 in CD_3OD. * The peaks due to a minor contaminant $PhCH_2OH$ are indicated by an asterisk.



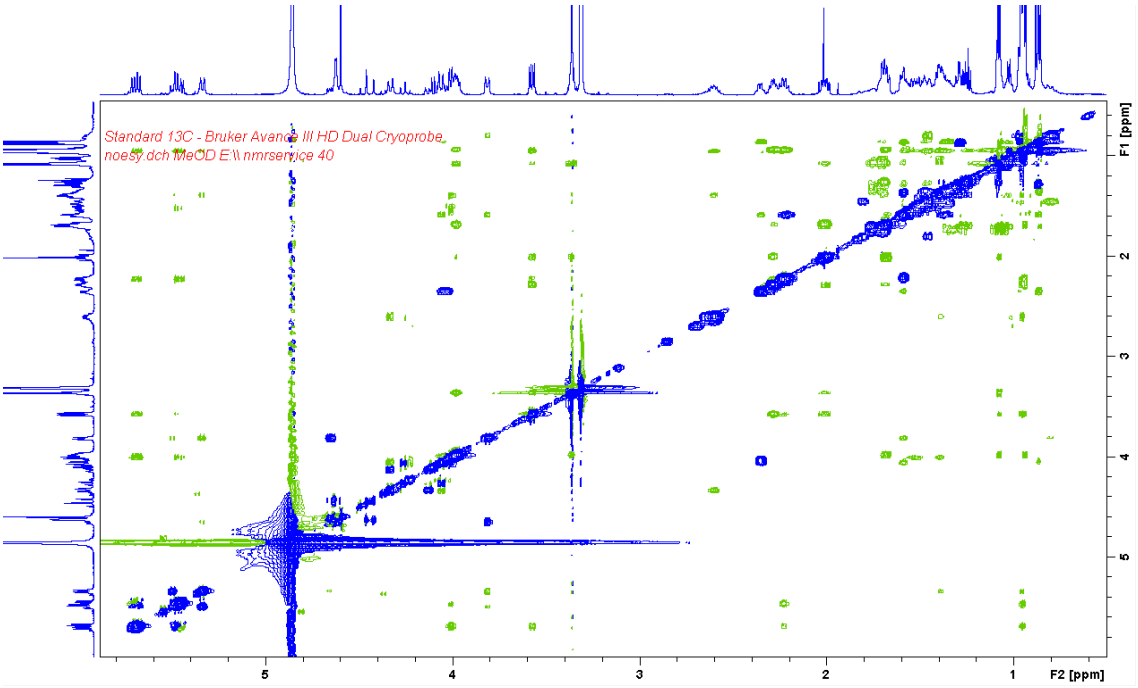
Appendix Figure 18
HSQC spectrum of intermediate T-22 in CD_3OD. Green peaks are CH_2 groups, and blue peaks are CH or CH_3.



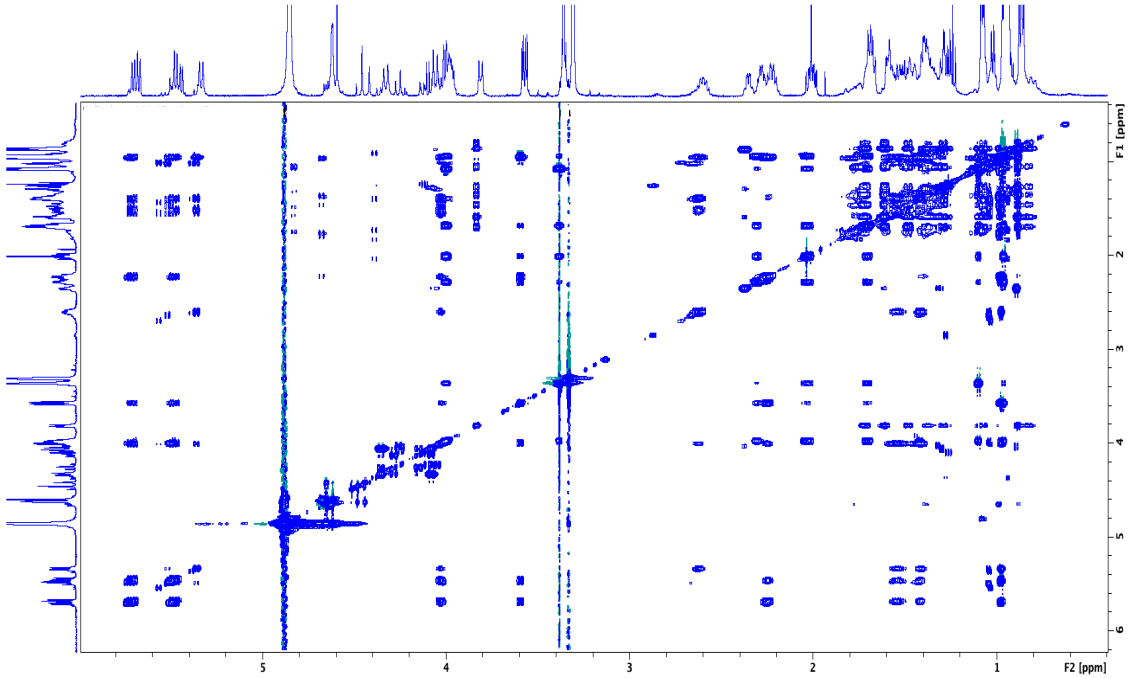
Appendix Figure 19
HMBC spectrum of T-22 in CD_3OD.



Appendix Figure 20
^{1}H-^{1}H COSY spectrum of T-22 in CD_3OD.



Appendix Figure 21
^{1}H-^{1}H NOESY spectrum of T-22 in CD_3OD. Green off-diagonal peaks show positive NOEs. Blue off-diagonal peaks (negative) show saturation transfer between equivalent protons in the interconverting isomers.



Appendix Figure 22
^{1}H-^{1}H TOCSY spectrum of T-22 in CD_3OD.

Appendix Tables

Appendix Table 1

NMR spectroscopic data for intermediate T-16 (Table prepared by Dr. Finian Leeper)

^{13}C ppm[a]	^{1}H (HSQC) ppm[b]	^{1}H (HMBC) ppm[b]	Assignment	COSY[c] and/or NOESY[d]
157.07	7.12 (1H, br d, *J* 8.3 Hz)	1.91, 1.063	C-5	2.64, 1.91, 1.46[d], 1.063[d]
140.1	5.23 (1H, d, *J* 10.0 Hz)	5.91, 4.29 & 4.25, 0.96	C-13	5.91[c], 2.56, 0.96[d]
136.12	5.71 (1H, dd, *J* 15.0 & 7.1 Hz)	4.03, 3.51, 2.22, 0.917	C-19	5.49, 4.03[d], 3.51[d], 2.22[c], 0.917[d]
135.88	-	5.77, 4.29 & 4.25	C-12	
132.37	5.91 (1H, d, *J* 15.8 Hz)	5.23, 4.29 & 4.25, 2.09	C-11	5.77, 5.23[c], 4.29[d] & 4.25[d], 2.09[c]
131.96	5.49 (1H, dd, *J* 15.0 & 7.2 Hz)	4.03, 2.22, 1.52-1.42	C-18	5.71, 4.03[c], 2.22[c], 0.917[d]
128.73	5.77 (1H, dt, *J* 15.8 & 7.0 Hz)	5.91, 2.09	C-10	5.91, 4.29[d] & 4.25[d], 2.09[c]
-	4.59 (2H, br s)	-	34-CH_2	-
86.16	3.51 (1H, dd, *J* 9.6 & 5.0 Hz)	5.71, 3.97, 2.22, 1.64, 0.917	C-21	5.71[d], 5.49[d], 2.22[c], 1.96[d], 0.917[d]
80.32	3.97 (1H, ddd, *J* 8.8, 6.7, 4.6 Hz)	3.32, 1.96, 1.056	C-24	3.32[c], 1.96 & 1.64[c], 1.056[d], 0.902[d]
79.68	3.32 (1H, qd, *J* 6.1 & 4.6 Hz)	3.35, 1.96, 1.056	C-25	3.97[c], 1.056[c]
73.52	4.03 (1H, q, *J* 6.5 Hz)	5.71, 5.49, 1.50 & 1.44	C-17	5.71[d], 5.49[c], 1.50[c] & 1.44[c]
57.82	4.29 & 4.25 (2 x 1H, d, *J* 11.7 Hz)	5.91, 5.23	C-30	5.91[d], 5.77[d], 2.56[d]
57.57	3.35 (3H, s)	3.32	OMe	1.056[d]
37.08	2.22 (1H, m)	5.71, 5.49, 3.51, 0.917	C-20	5.71[c], 5.49[c], 3.51[c], 0.917
36.23	1.46 (2H, m)	2.09, 1.42, 1.063	C-7	7.12[d], 2.64[c], 1.91[d], 1063[d]
35.46	1.96 (1H, m)	3.32, 2.24, 0.902	C-23	3.97[c], 3.51[d],3.32[d], 2.24[c], 1.64[c]
	& 1.64 (1H, dd, *J* 12.4, 6.9 Hz)			3.97[c], 1.96[c], 0.902[d]
35.33	2.24 (1H, m)	1.96 & 1.64, 0.902	C-22	3.51[d], 1.96, 0.902[c]
35.01	1.50 & 1.44 (each 1H, m)	5.49, 4.03	C-16	4.03[c]
34.51	2.64 (1H, m)	1.46, 1.063	C-6	7.12, 1.91[d], 1.46[c], 1.063[c]
33.35	1.36 & 1.30 (each 1H, m)	5.23, 4.03, 1.50, 0.96	C-15	2.56
33.24	2.09 (2H, m)	5.91, 5.77, 1.47	C-9	5.91[c], 5.77[c], 1.42[c]
32.62	2.56 (1H, m)	1.50-1.35, 0.96	C-14	5.23[c], 4.29[d] & 4.25[d], 1.36 & 1.30, 0.96[c]
27.5	1.42 (2H, m)	5.77, 2.09, 1.46	C-8	2.09[c]
21.78	0.96 (3H, d, *J* 6.6 Hz)	5.23, 2.56, 1.36 & 1.30	C-29	5.23[d], 2.56[c]
19.82	1.063 (3H, d, *J* 6.6 Hz)	2.64, 1.91, 1.46	C-31	7.12[d], 2.64[c]
17.1	0.917 (3H, d, *J* 6.8 Hz)	5.71, 3.51, 2.22	C-28	5.71[d], 5.49[d], 3.51[d], 2.22
15.88	1.056 (3H, d, *J* 6.3 Hz)	3.97, 3.32	C-26	3.97[d], 3.32[c], 1.96[d]
13.79	0.902 (3H, d, *J* 7.0 Hz)	3.51, 1.96 & 1.64	C-27	3.97[d], 2.24c, 1.64[d]
12.32	1.91 (3H, s)	7.12 (w)	C-32	7.12, 2.64[d]

[a]relative to ^{13}C in $CDCl_3$ at 77.23. [b]relative to residual protons in $CDCl_3$ at 7.24. [c]COSY and NOESY, [d]NOESY only, otherwise COSY only.

Appendix Table 2

NMR spectroscopic data for intermediate T-22 (Table prepared by Dr. Finian Leeper)

^{13}C ppm[a]	^{1}H (HSQC) ppm[b]	^{1}H (HMBC) ppm[b]	Assignment	COSY[c] and/or NOESY[d]
179.99 v.br	-	4.62	} C-33	-
174.72 br	-	4.62	} C-1	-
142.76 br	5.34 (1H, d, *J* 9.7 Hz)	4.33 & 4.06, 3.81, 0.952	C-13	2.60[c], 3.81[d], 1.40[d], 0.952[d]
136.17	5.69 (1H, dd, *J* 15.6 & 7.5 Hz)	4.01, 3.57, 2.225, 0.954	C-19	5.46, 2.225[c], 4.01[d], 3.57[d], 0.954[d]
135.18 br	-	4.06, 3.81	C-12	-
133.47	5.46 (1H, dd, *J* 15.6 & 7.5 Hz)	4.01, 2.225, 1.52	C-18	5.69, 4.01[c], 2.225[d], 1.52[d], 0.954[d]
104.50 br	-	0.873	C-3	-
99.38 v.br	-	4.62	C-2	-
87.8	3.57 (1H, dd, *J* 9.2 & 4.2 Hz)	5.69, 3.96, 2.225, 1.68, 0.940	C-21	2.285[c], 2.225[c], 5.69[d], 3.36[d], 2.01[d], 0.954[d]
83.57	3.81 (1H, br d, *J* 9.7 Hz)	5.34, 4.33 & 4.06	C-11	1.59[c], 5.34[d]
81.54	3.98 (1H, m) (ddd?)	3.36, 201, 1.076	C-24	3.36[c],2.01[c], 1.68[c], 1.076[d], 0.940[d]
80.85	3.36 (1H, m)	3.36, 201, 1.076	C-25	3.98[c], 1.076[c], 3.57[d], 2.01[d]
73.9	4.01 (1H, m) (d?)	5.69, 5.46, 1.52	C-17	5.46[c], 1.52[c], 5.69[d], 1.40[d]
68.32	4.62 (2H, br s)	-	C-34	-
57.76	3.36 (3H, s)	3.36	OMe	-
57.17	4.33 & 4.06 (each 1H, br d, *J* 11.0 Hz)	5.34, 3.81	C-30	[e] 2.60[d], 1.59[d]
44.98	1.59 (1H, m)	3.81, 2.34, 1.70, 1.59, 0.862	C-5	2.355[c], 1.35
38.59	2.225 (1H, dt, *J* 9.0 & 7.0 Hz)	5.69, 5.46, 0.954	C-20	5.69[c], 3.57[c], 0.954[c], 5.46[d]
37.01	1.70 & 1.06 (each 1H, m)	0.862	C-7	[e] 1.27[d]
36.52	2.285 (1H, br q, *J* 6.0 Hz)	2.01 & 1.68, 0.940	C-22	3.57[c], 2.01[c] & 1.68[c], 0.940[c]
36.43	1.59 (1H, m)	3.81, 2.355, 1.59	C-10	3.81[c], 0.80, 4.06[d]
36.20*	} 2.01 (1H, ddd, *J* 12.5, 8.6 & 6.6 } Hz) & 1.68 (1H, m)	} 3.57, 3.36, 2.285, } 0.940	} C-23	[e] 3.98[c], 2.285[c], 3.57[d], 3.36[d], 1.076[d]
36.18*	} 1.52 (2H, m)	} 5.46, 2.60	} C-16	4.01[c], 1.40, 5.46[d]
35.51	2.355 (1H, qd, *J* 7.2 & 2.2 Hz)	0.873	C-4	1.59[c], 0.873[c]
34.53	1.35 (1H, m)	1.69, 1.59, 0.862	C-6	1.59, 0.862[c]
34.11	1.40 (2H, m)	5.34, 4.01, 1.52, 0.952	C-15	2.60[c], 1.52, 5.34[d], 4.01[d], 0.952[d]
33.19	2.60 (1H, m)	5.34, 1.52, 0.952	C-14	5.34, 1.40[c], 0.952[c], 4.33[d]
30.15	1.46 (br d, *J* 12 Hz) & 0.80 (br q, *J* 12 Hz)	3.81	C-9	[e] 1.58
26.56	1.69 & 1.27 (each 1H, m)	-	C-8	[e] 1.06, 1.70[d]
22.01	0.952* (3H, d, *J* 6.6 Hz)*	5.34, 2.60, 1.40	C-29	2.60, 5.34[d]
19.06	0.862 (3H, d, *J* 6.2 Hz)	1.06	C-31	1.35[c], 2.355[d]
17.59	0.954* (3H, d, *J* 6.6 Hz)*	5.69, 3.57, 2.225	C-28	2.225[c], 5.69[d], 5.46[d]
15.87	1.076 (3H, d, *J* 6.3 Hz)	3.98, 3.36	C-26	3.36[c], 3.98[d], 2.01[d]
14.04	0.940 (3H, d, *J* 6.8 Hz)	3.57, 2.28, 2.01, 1.68	C-27	2.285[c], 3.98[d], 2.225[d], 1.68[d]
10.08	0.873 (3H, d, *J* 6.6 Hz)	2.355, 1.59	C-32	2.355[c], 1.59[d]

[a]relative to ^{13}C in CD_3OD at 49.15. [b]relative to residual protons in CD_3OD at 3.31. [c]COSY and NOESY, [d]NOESY only, otherwise COSY only. [e]these diastereotopic protons showed COSY & NOESY correlations with each other.*Signals too close to be distinguished in the 2D NMR spectra are indicated by asterisks

Appendix table 3

Crystallographic data (Table prepared by Fernanda Paiva). Values in parentheses are for the high resolution outer shell.

PDB entry	Tsn15 (6NOI)	Tsn15-ligand (6NNW)
Beamline	Diamond Light Source (DLS) I04 beamline, Oxford, UK	PETRAIII, beamline 13, Hamburg, Germany
Resolution range (Å)	29.7 - 1.80 (1.84 - 1.80)	43.81 - 1.7 (1.76 - 1.70)
Space group	C 1 2 1	C 1 2 1
Unit cell (Å)	128.8 36.9 91.8 90 107.5 90	129.7 36.9 92.0 90 107.8 90
Total reflections	239480 (14251)	90988 (8254)
Unique reflections	38365 (2014)	45893 (4205)
Multiplicity	6.2 (3.9)	2.0 (2.0)
Completeness (%)	99.0 (88.7)	98.8 (91.1)
Mean I/sigma(I)	23.90 (6.56)	13.3 (2.0)
Wilson B-factor	15.47	21.87
R-merge	0.063 (0.218)	0.030 (0.338)
R-meas	0.068 (0.253)	0.042 (0.479)
R-pim	0.027 (0.125)	0.0298 (0.339)
CC1/2	0.998 (0.949)	0.998 (0.788)
CC*	1 (0.888)	1 (0.939)
Reflections used in refinement	38363 (3503)	45883 (4202)
Reflections used for R-free	1837 (173)	2318 (195)
R-work	0.156 (0.191)	0.170 (0.299)
R-free	0.186 (0.233)	0.204 (0.355)
CC(work)	0.969 (0.932)	0.968 (0.867)
CC(free)	0.967 (0.896)	0.961 (0.812)
Number of non-hydrogen atoms	3669	3562
macromolecules	3228	3211
ligands	34	82
solvent	407	269
Protein residues	412	413
RMS(bonds) (Å)	0.006	0.007
RMS(angles) (Å)	0.89	0.9
Ramachandran favored (%)	99	99.51
Ramachandran allowed (%)	0.7	0.49
Ramachandran outliers (%)	0	0
Rotamer outliers (%)	0.56	0
Clashscore	5.98	4.16
Average B-factor (Å^2)	22.8	28.97
macromolecules	21.04	27.84
ligands	37.10	39.71
solvent	35.04	39.2
Number of TLS groups	1	1